
DIPLOMARBEIT

Herr
Michael Konietzky

**Entwicklung eines forensischen
Analysewerkzeuges zur
Untersuchung von Nachrichten
in sozialen Netzwerken am
Beispiel von Twitter**

Mittweida, 2014

DIPLOMARBEIT

Entwicklung eines forensischen Analysewerkzeuges zur Untersuchung von Nachrichten in sozialen Netzwerken am Beispiel von Twitter

Autor:

Herr Michael Konietzky

Studiengang:

Multimediatechnik

Seminargruppe:

MK06w2

Erstprüfer:

Prof. Dr. rer. nat. Dirk Labudde

Zweitprüfer:

M.Sc. Michael Spranger

Einreichung:

17.April 2014

Verteidigung/Bewertung:

Mittweida, 2014

DIPLOMTHESIS

Development of a forensic analysis tool to study news in social networks using the example of Twitter

author:

Mr. Michael Konietzky

course of studies:

Multimediatechnik

seminar group:

MK06w2

first examiner:

Prof. Dr. rer. nat. Dirk Labudde

second examiner:

M.Sc. Michael Spranger

submission:

17.April 2014

defence/evaluation:

Mittweida, 2014

Bibliografische Beschreibung:

Konietzky, Michael:

Entwicklung eines forensischen Analysewerkzeuges zur Untersuchung von Nachrichten in sozialen Netzwerken am Beispiel von Twitter. - 2014. - 12 Seiten Verzeichnisse, 57 Seiten Inhalt, 1 Seite Anhänge S. Mittweida, Hochschule Mittweida, Fakultät Elektro- und Informationstechnik, Diplomarbeit, 2014

Referat:

Die vorliegende Arbeit befasst sich mit der prototypischen Entwicklung eines Analysewerkzeuges zur Unterstützung von forensischen Untersuchungen im Bereich sozialer Netzwerke, konkret am Beispiel von Twitter. Für die kriminalistischen Ermittler wird eine Analysesoftware konzipiert, die Nachrichten des Mikrobloggeringsystems Twitter automatisch nach fallrelevanten Informationen untersucht.

Danksagung

Besonderer Dank gilt Herrn M.Sc. Michael Spranger und Herrn Professor Dr. rer. nat. Dirk Labudde für die Bereitstellung dieses aktuellen und interessanten Themas. Speziell Herrn M.Sc. Michael Spranger möchte ich nochmals explizit meinen Dank für die hervorragende Betreuung und Geduld, die er mir während der Entstehungsphase dieser wissenschaftlichen Arbeit entgegengebracht hat, ausrichten.

Des Weiteren gilt mein Dank Frau Luisa Appelt, die ebenfalls an diesem Projekt mitgewirkt hat und mir ebenso bei vielen Fragen tatkräftig zur Seite stand.

Vielen Dank auch an Frau Anne-Sophie Schultka, die mir sehr bei der Gestaltung sowie der Grammatik und Orthografie geholfen hat.

Inhaltsverzeichnis

1	Vorwort	13
1.1	Hintergrund der Arbeit	13
1.2	Aufgabenstellung	14
1.3	Aufbau der Arbeit	15
2	Entwicklung der computervermittelten Kommunikation im Internet	16
2.1	Grundlagen	16
2.1.1	Historische Entwicklung des Internets	16
2.1.2	Rechnernetzwerke	17
2.1.2.1	Netzwerktopologien	18
2.1.2.1.1	Bustopologie	18
2.1.2.1.2	Ringtopologie	18
2.1.2.1.3	Sterntopologie	19
2.1.2.2	Lokale Netze (LANs)	20
2.1.2.2.1	Ethernet	20
2.1.2.2.2	WLAN	21
2.1.2.3	Weitverkehrsnetz	22
2.2	Computervermittelte Kommunikationsformen	23
2.2.1	Asynchrone computervermittelte Kommunikation	23
2.2.1.1	E-Mail	23
2.2.1.2	Mailinglisten	25
2.2.1.3	Newsgroups	25
2.2.1.4	Internetforen	26
2.2.1.5	Blog	26
2.2.2	Synchrone computervermittelte Kommunikation	27
2.2.2.1	Internet-Telefonie	27
2.2.2.2	Chats	29
2.2.2.3	MUDs und MOOS	30
2.2.3	Soziale Netzwerke und Personal Publishing	31
3	Der Microbloggingdienst Twitter	32
3.1	Microblogging	32
3.2	Twitter	32
3.2.1	Anmeldung bei Twitter	34

3.2.2	Twitter-Suchmaschine	35
3.2.3	Zugriffsmöglichkeiten auf Twitter	36
3.2.3.1	Twitter-API	36
3.2.3.2	Echtzeit-API Firehose	38
3.2.3.3	Zugriffe über Reseller	39
3.2.4	Charakteristika von Tweets	40
4	Forensische Ansätze	43
4.1	Forensische Informatik	43
4.2	Kriminalität in sozialen Netzwerken	43
4.3	Computerlinguistische Analyse von Twitterdaten	44
4.4	Emotionserkennung in sozialen Netzwerken	45
4.5	Softwarebasierte und analytische forensische Ansätze zur Personensuche in sozialen Netzwerken	47
4.5.1	First Story Detection (FSD)	47
4.5.2	soziale netzwerkbasierende Entitätsgewinnung für Personenontologien	48
4.5.3	Erzeugung von Bekanntenempfehlungen in sozialen Netzwerken durch Mining des Benutzercontent	49
5	TwitterSearcher	51
5.1	Theoretische Grundlagen	51
5.1.1	Eclipse	51
5.1.2	OSGi-Service Plattform	53
5.1.3	Rich Client Plattform	56
5.1.4	Eclipse Modeling Framework	57
5.2	Programmentwicklung des TwitterSearchers	57
5.2.1	Motivation	57
5.2.2	Use Case Diagramm	58
5.2.3	Einbinden der Twitter-API	60
5.2.4	Entwurf der grafischen Benutzeroberfläche	61
5.2.4.1	Window Builder	61
5.2.4.2	Grafische Benutzeroberfläche des TwitterSearcher	61
5.2.4.2.1	Details Part	63
5.2.4.2.2	Query Overview Part	66
5.2.5	Datenmodell	67
5.2.6	Services	68

5.2.7	Handlerklassen	69
6	Schlussbetrachtung	70
	Literaturverzeichnis	
	Anhang	

Abbildungsverzeichnis

Abbildung 1 Langfristige Kriminalitätsentwicklung in Deutschland [PKS12]	13
Abbildung 2 Cybercrime im engeren Sinne [BKA12]	14
Abbildung 3 Zeitleiste: Entwicklung des Internets und ausgewählte Anwendungen [SCH07]	17
Abbildung 4 Schema einer Bustopologie [BOD06]	18
Abbildung 5 Schema einer Ringtopologie [BOD06]	19
Abbildung 6 Schema einer Sterntopologie [BOD06]	19
Abbildung 7 Schema WLAN [WLA14]	21
Abbildung 8 Kommunikationsprinzip Bluetooth [KIO14]	22
Abbildung 9 Schema: prinzipieller Ablauf einer E-Mail Zustellung [WIK143]	24
Abbildung 10 Funktionsprinzip Internet Telefonie [MOO14]	28
Abbildung 11 Oberfläche der Twitter-Suche	35
Abbildung 12 schemenhafte Funktionsweise von DataSift [DAT14]	40
Abbildung 13 Beispieltweet mit den wichtigsten Bestandteilen	42
Abbildung 14 Emotionsontologie für die 6 Ekman's Basisgefühle + Liebe [ROB12]	46
Abbildung 15 Beispiel einer einfachen Ontologie [HOL13]	49
Abbildung 16 Übersicht der Eclipse-Versionen [WIK144]	51
Abbildung 17 Plattformarchitektur von Eclipse [ECL14]	52
Abbildung 18 Eclipse Architektur Version 4 [VOG13]	53
Abbildung 19 OSGi-Framework Architektur [OSG14]	54
Abbildung 20 gemeinsame Nutzung eines Services von zwei Bundles [OSG14]	55
Abbildung 21 Eclipse RCP-Architektur [VOG13]	56
Abbildung 22 Use Case-Diagramm für den TwitterSearcher	58
Abbildung 23 Ausschnitt aus der Ordnerstruktur des TwitterSearchers mit Referenced Libraries	60
Abbildung 24 Ausschnitt der Windowstruktur aus der Application.e4xmi	62
Abbildung 25 Windowstruktur des Details Part	63
Abbildung 26 Searchform des Detail Parts	64
Abbildung 27 Result- und Selected Tweetform des Detail Parts	65
Abbildung 28 grafische Benutzeroberfläche des TwitterSearchers mit gekennzeichnetem Query Overview Part	66
Abbildung 29 EMF-Datenmodell des TwitterSearchers	67

Abkürzungsverzeichnis

BAK	Bundeskriminalamt
ENIAC	Electronic Numerical Integrator and Computer
ARPANET	Advanced Research Projects Agency Network
USENET	Unix User Network
TCP	Transmission Control Protocol
FTP	File Transfer Protocol
Telnet	Telecommunication Network
SMTP	Simple Mail Transfer Protocol,
POP	Post Office Protocol,
IMAP	Internet Message Access Protocol,
WWW	World Wide Web,
IRC	Internet Relay Chat,
UUCP	Unix to Unix CoPy,
NNTP	NetNews Transfer Protocol,
IEEE	Institute of Electrical and Electronics Engineers,
WLAN	Wireless Local Area Network
DSL	Direct Subscriber Line
DPU	Data Processing Unit
CSDL	Curated Stream Definition Language
IDE	Integrated Development Environment
SDK	Software Development Kit
SWT	Standard Widgeting Toolkit
OSGi	Open Service Gateway initiative
JRE	Java Runtime Environment
UML	Unified Modeling Language
XML	Extensible Markup Language
XMI	XML Metadata Interchange

1 Vorwort

1.1 Hintergrund der Arbeit

Die Kommunikation in sozialen Netzwerken und Nachrichtendiensten erlebte in den letzten Jahren einen ständigen Anstieg. Viele Nutzer teilen über diverse Plattformen persönliche Meinungen, Gedanken oder auch aktuelle Tätigkeiten mit. Nicht selten kommt es dabei vor, dass Täter Straftaten unter Nutzung derartiger Medien organisieren, in dem Glauben der eigenen Anonymität.

Während bei klassischen Delikten, wie Rauschgiftvergehen, Körperverletzungen, Sachbeschädigungen, Sexualverbrechen, schwerem beziehungsweise leichtem Diebstahl und Betrug, das Niveau relativ gleichbleibend ist, wie die polizeiliche Kriminalstatistik des BKA belegt [PKS12], ist ein Anstieg der Cyberkriminalität zu verzeichnen. Cyberkriminalität im engeren Sinne beschreibt hierbei „alle Straftaten, die unter Ausnutzung moderner Informations- und Kommunikationstechnik, auch gegen diese selbst begangen wurde“ [BKA12 Seite 3]. Der Freistaat Sachsen belegt in der Statistik der Bundesländer den 7.Rang mit einem Strafanteil von 5,2% und befindet sich damit im unteren Mittelfeld [PKS12].

Jahr	Einwohner am 01.01. *)	Fälle	Veränderung gegenüber dem Vorjahr in %	Häufig- keitszahl **)	Veränderung gegenüber dem Vorjahr in %	Aufklärung in %
1998	82.057.400	6.456.996	-2,0	7.869	-2,0	52,3
1999	82.037.000	6.302.316	-2,4	7.682	-2,4	52,8
2000	82.163.500	6.264.723	-0,6	7.625	-0,7	53,2
2001	82.259.500	6.363.865	1,6	7.736	1,5	53,1
2002	82.440.300	6.507.394	2,3	7.893	2,0	52,6
2003	82.536.700	6.572.135	1,0	7.963	0,9	53,1
2004	82.531.700	6.633.156	0,9	8.037	0,9	54,2
2005	82.501.000	6.391.715	-3,6	7.747	-3,6	55,0
2006	82.438.000	6.304.223	-1,4	7.647	-1,3	55,4
2007	82.314.900	6.284.661	-0,3	7.635	-0,2	55,0
2008	82.217.800	6.114.128	-2,7	7.436	-2,6	54,8
2009	82.002.400	6.054.330	-1,0	7.383	-0,7	55,6
2010	81.802.300	5.933.278	-2,0	7.253	-1,8	56,0
2011	81.751.602	5.990.679	1,0	7.328	1,0	54,7
2012	81.843.743	5.997.040	0,1	7.327	0,0	54,4

Abbildung 1 Langfristige Kriminalitätsentwicklung in Deutschland [PKS12]

*) Quelle statistisches Bundesamt

**) Häufigkeitszahl: Fälle pro 100000 Einwohner

Während Täter das Internet und dessen Dienste für Straftaten beziehungsweise die Organisation solcher nutzen, müssen die Ermittler zur Aufklärung ihrer Fälle den Tätern mit gleicher Technik entgegentreten.

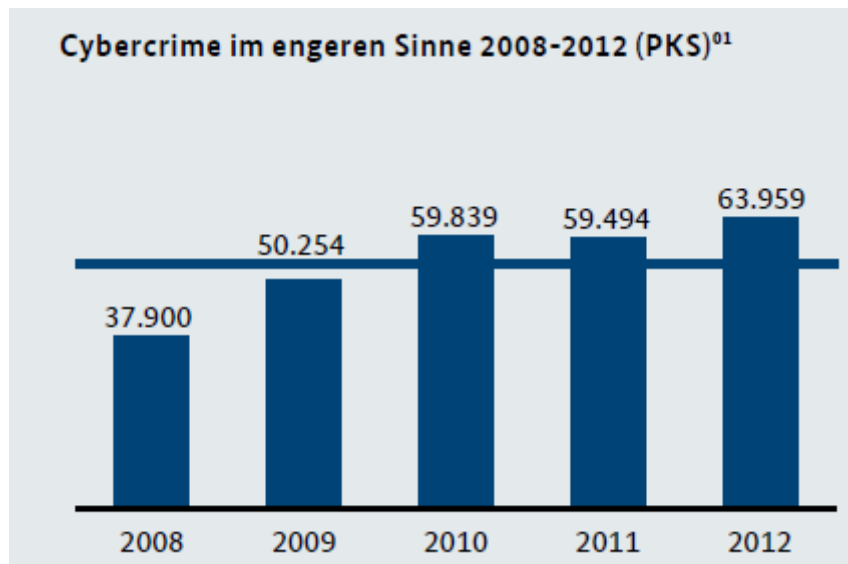


Abbildung 2 Cybercrime im engeren Sinne [BKA12]

Stellt man sich zum Beispiel eine Brandserie vor, die noch keinen Ermittlungserfolg liefern konnte, können die kriminalistischen Ermittler auf Grundlage des Wissens, dass soziale Netzwerke als Kommunikationsmittel für Straftaten und deren Organisation verwendet werden, in diesem Umfeld nach Hinweisen auf Täter, Opfer oder Zeugen nachgehen. Weil die manuelle Suche nach relevanten Texten und Mitteilungen sehr zeitaufwändig ist, wäre die Entwicklung einer Software hilfreich, die diese Aufgabe für die kriminalistischen Ermittler übernehmen könnte.

1.2 Aufgabenstellung

Das Projekt eines forensischen Analysewerkzeuges zur Untersuchung von sozialen Netzwerken basiert auf einer Kooperation zwischen der Polizeidirektion Chemnitz und der Hochschule Mittweida. Im Einzelnen soll dabei für die kriminalistischen Ermittler eine Möglichkeit geschaffen werden, in öffentlichen Bereichen sozialer Netzwerke nach Hinweisen zu ungeklärten und geplanten Straftaten zu fahnden und somit die forensischen Untersuchungen zu unterstützen.

Einer der bekanntesten öffentlichen Nachrichtendienste ist dabei Twitter. Die Faktoren, dass für den Microbloggingdienst eine frei zugängliche Programmierschnittstelle, eine sogenannte API, existiert die es dem Programmierer ermöglicht alternative Programmbibliotheken in den eigenen entwickelten Programmen zu nutzen, ohne den Code des Anwendungsprogrammes ändern zu müssen¹ und der Umstand der frei zugänglichen Nachrichten ermöglichen eine rechtlich problemlose Analyse.

Das Ergebnis dieser wissenschaftlichen Arbeit ist die Entwicklung eines Analysewerkzeuges, das alle Nachrichten (Tweets) von Twitter nach vordefinierten Begriffen untersucht und sowohl die Eingabe als auch die Ausgabe strukturiert darstellt. Die gefundenen Suchergebnisse werden dabei mit computerlinguistischen Methoden auf ihre Relevanz

untersucht. Erschwert wird die Suche dabei durch die Charakteristik der Texte, da mittlerweile viele umgangssprachliche Begriffe und Abkürzungen in den Nachrichten enthalten sind. Des Weiteren bestehen für kriminalistischen Ermittler die Möglichkeiten einer Volltext- beziehungsweise Einzelwörtersuche sowie die Funktionalität der Fremdsprachensuche.

1.3 Aufbau der Arbeit

Kapitel 1 wird als Einführung die historische Entwicklung der Kommunikation im Internet, vom einfachen Chat hin zu sozialen Netzwerken, näher erläutern.

Kapitel 2 beginnt mit einer kontroversen Betrachtung sozialer Netzwerke und diskutiert die Entscheidungsgründe für Twitter und warum die Wahl nicht auf Facebook oder lokale Netzwerke fiel. Weiterführend in den Unterkapiteln erfolgt eine Erläuterung zu dem Begriff Mikrobloggingdienst und welche Alternativen zu Twitter bestehen. Das Kapitel wird abgeschlossen mit Themen zur Erläuterungen der Twitteranmeldung, Twitter-API und der Nachrichtenform Tweet. In Bezug auf Tweets untersucht das entsprechende Kapitel die Charakteristik einer solchen Nachricht unter anderen auf Schreibfehler, Hashtagunterbrechungen oder Emotionserkennung und wie sich diese Charakteristika hinsichtlich der forensischen Untersuchung auswirken.

Kapitel 3 befasst sich mit dem Thema Forensik. In den Unterkapiteln wird dabei auf die Methoden der forensischen Computerlinguistik eingegangen. Hierbei wird untersucht, in welcher Weise Nachrichten die forensischen Ermittlungen unterstützen können anhand verschiedener forensischer Softwaremodelle und welche Hinweise auf Straftaten in den Nachrichten enthalten sind unter Berücksichtigung der oben geschilderten Charakteristika.

Im 4. Kapitel wird die Funktionsweise des Twitter Searchers näher erläutert. Beginnend mit der Entwicklungsumgebung Eclipse und der Programmiersprache Java, in der das forensische Analysewerkzeug entwickelt wird. Überleitend erfolgt eine Betrachtung hinsichtlich der Funktionsweise, Erläuterungen zum Quelltext, Einbindung der Twitter-API, Oberflächenentwurf, Anforderungsbetrachtung in Form eines Use Case Diagramms sowie die Verwendung des WindowBuilder in Eclipse.

Das 5. Kapitel bildet mit der abschließenden Einschätzung hinsichtlich der Probleme, die während der Entwicklungsphase aufgetreten sind und deren Lösung, sowie einen Ausblick bezüglich der Erweiterungsmöglichkeiten des Analysewerkzeuges den Abschluss dieser wissenschaftlichen Arbeit.

2 Entwicklung der computervermittelten Kommunikation im Internet

2.1 Grundlagen

2.1.1 Historische Entwicklung des Internets

Wenn in der Vergangenheit nicht einige Pioniere mit ihren Erfindungen dafür gesorgt hätten, die moderne Informatik zu bereichern, wäre die Nutzung des Internets sowie seiner Kommunikationsdienste, zu denen heute auch soziale Netzwerke zählen, nicht möglich. Die historische Entwicklung der Kommunikation im Internet ist stark verbunden mit der Erfindung des modernen Computers. Als Vorlage moderner Digitalrechner gilt die von Charles Babbage 1837 entwickelte Rechenmaschine „Analytical Engine“. Den ersten funktionsfähigen Digitalrechner (Z3) entwarf Konrad Zuse 1941 in Berlin. John Presper Eckert und John William Mauchly entwickelten fünf Jahre später an der University of Pennsylvania den ENIAC, den ersten rein elektronischen Universalrechner, der für die Berechnung ballistischer Tabellen für die US-Armee fungierte [WDR09].

Mitte der 60er Jahre wurden in den USA erste Versuche unternommen, mehrere Computer miteinander zu verbinden. Diese Forschungsprojekte wurden vom Militär finanziert um eine gleichmäßigere Auslastung von Rechenkapazitäten und damit Kostenersparnisse zu erzielen [FRA12]. 1969 wurde das Arpanet, welches ursprünglich von der US-Luftwaffe 1962 in Auftrag gegeben wurde um US-amerikanische Universitäten zu verbinden, die für das Verteidigungsministerium forschten, fertiggestellt. Der erste Netzknoten ging an der University of California in Betrieb und wenig später wurden die Universitäten von Utah, Stanford, Santa Barbara und Los Angeles miteinander vernetzt [BEC06]. Dabei wurde ein dezentrales System entwickelt, welches bei Ausfall von Teilen des Netzes nicht komplett zusammenbrechen würde. Die Datenströme wurden in kleinere Pakete zerlegt und beim Empfängerrechner wieder zusammengesetzt (paketvermittelt) [FRA12]. Mitte der 70er Jahre wurde das TCP im Arpanet entwickelt, dass die Grundlage des heutigen Internets darstellen sollte. Das TCP ist ein Protokoll, welches die Art und Weise regelt, wie Daten zwischen Rechnern ausgetauscht werden. Diverse Dienste wie FTP (Netzwerkprotokoll zur Übertragung von Dateien über IP-Adressen), Telnet (Netzwerkprotokoll für den Fernzugriff auf Rechnerkapazitäten) und E-Mail wurden bereits Anfang der 70er entworfen [BEC06]. Das Usenet stellte Ende der 70er Jahre eine öffentlich zugängliche Alternative zum Arpanet dar. Zunächst bestand nur die Möglichkeit in Diskussionsforen Probleme der UNIX-Anwender zu diskutieren [BEC06]. In den 80er kam es zum Aufbau weiterer staatlicher (für die US-Weltraumbehörde NASA) beziehungsweise kommerzieller Netze, wie zum Beispiel Telnet oder Corporate Networks. Im Herbst 1983 verband das Arpanet bereits 500 Rechner miteinander. Da den Universitäten das Arpanet nicht generell offen stand, begannen diese eigene Netzwerke wie CSNet oder BITNet zu entwickeln. 1990 wurde das Arpanet

abgeschaltet, da nun private Unternehmen die Gestaltung des Netzes übernahmen [BEC06]. 1989 entwickelten Tim Berners-Lee und Robert Cailliau auf Grundlage ähnlicher Technologien am Genfer Kernforschungszentrum CERN die Idee eines Hypertextprojektes mit dem Namen WorldWideWeb, welches auf dem Hypertext Transfer Protocol (http) und der grafikfähigen Hypertext Markup Language (HTML) basierte. Anfänglich war es für die Kooperation zwischen den CERN-Mitarbeitern konzipiert worden, doch 1993 wurde es für die Öffentlichkeit freigegeben [WDR09]. „1995 existieren bereits 50.000 Host-Rechner, von denen HTML-Dokumente abgerufen werden konnten“ [FRA12 Seite 10]. 1998 hat sich die Zahl auf 2 Millionen erhöht und trug dabei zu einem großen Anstieg des globalen Datenaustausches bei. Mit Beginn des 21. Jahrhunderts kam es zu einer Erweiterung der Bandbreite, die es möglich machte immer größere Datenmengen zur Verfügung zu stellen. 2001 ging die Enzyklopädie Wikipedia online, die als wichtigste Nachschlagequelle weltweit gilt [FRA12]. Oft wird das WorldWideWeb mit dem Internet gleichgesetzt, was allerdings nur einen Internetdienst darstellt.

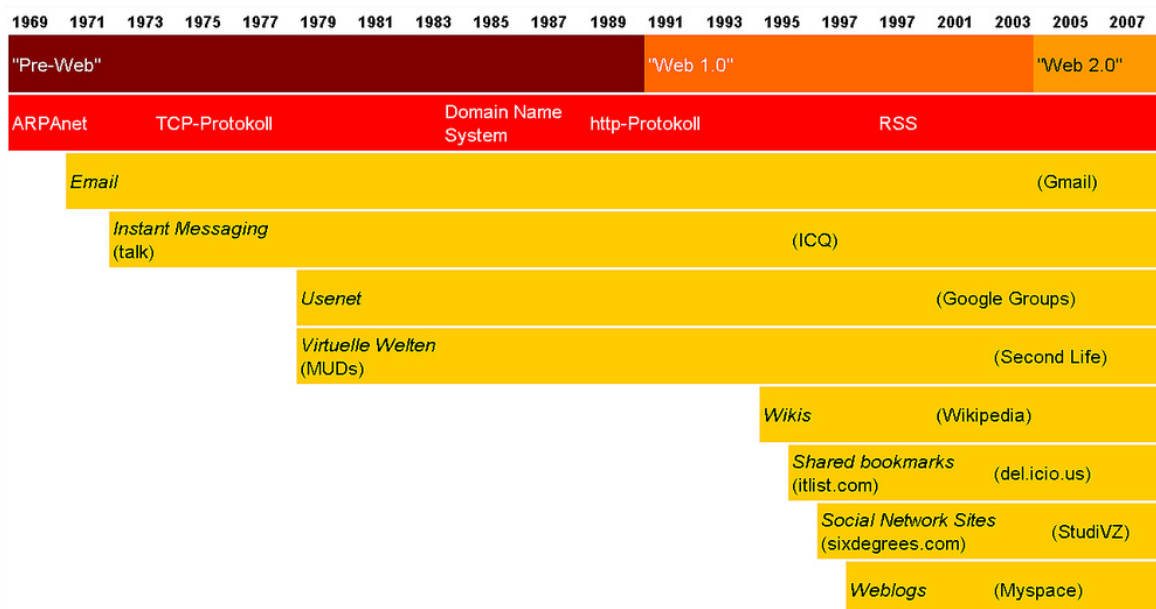


Abbildung 3 Zeitleiste: Entwicklung des Internets und ausgewählte Anwendungen [SCH07]

2.1.2 Rechnernetzwerke

Um Kommunikation im Internet betreiben zu können, müssen die Rechner der beteiligten Benutzer miteinander verbunden sein. Da Kommunikation mindestens zwei Teilnehmer voraussetzt besteht auch ein Rechnernetzwerk aus mindestens zwei PCs, einem Sender und einem Empfänger, die über ein Übertragungsmedium miteinander kommunizieren. Es ist jederzeit möglich einen weiteren PC in ein Netzwerk hinzuzufügen. Als Übertragungsmedium können ein Koaxialkabel, ein Twisted Pair Kabel, ein Glasfaserkabel oder diverse drahtlose Übertragungsformen (Funk, Mikrowellen, Infrarot, Lichtwellen u.a.) verwendet werden.

2.1.2.1 Netzwerktopologien

2.1.2.1.1 Bustopologie

Bei der Bustopologie erfolgt die Anordnung aller teilnehmenden Geräte mit Stichleitungen an einem zentralen Kabel. Dieser Aufbau ist dezentral, was bedeutet, dass alle Geräte gleichrangig am Bus angeschlossen sind. Der Nachteil dieser Anordnung besteht darin, dass eine Störung auf dem Hauptkabel auch alle angeschlossenen Geräte betrifft. Die Bustopologie erfährt aufgrund dieses Nachteils nur noch wenige Einsatzmöglichkeiten [LIN13].

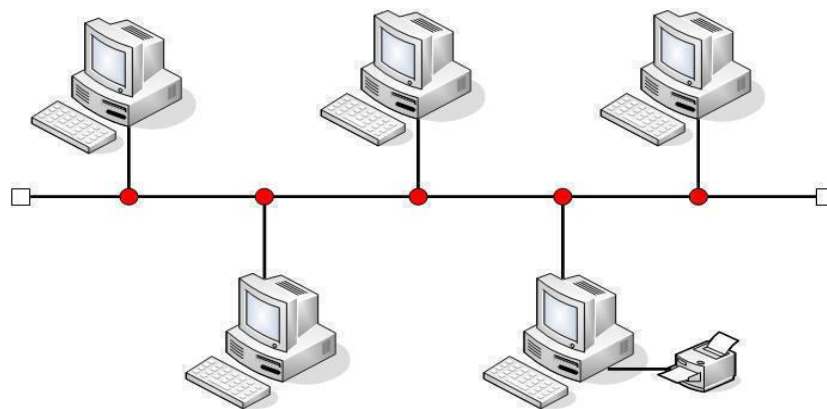


Abbildung 4 Schema einer Bustopologie [BOD06]

2.1.2.1.2 Ringtopologie

Ein Beispiel für eine Ringtopologie ist *Token-Ring*. Dabei wandert ein Token (spezielles Datenpaket) im Kreis. Möchte ein Gerät senden, muss es den Token besitzen, der zuvor von einer Monitorstation erzeugt wurde. Jedes Gerät darf den Token aber nur eine bestimmte Zeit behalten, damit jedes Gerät die Möglichkeit besitzt Daten zu versenden. Anschaulich gesprochen könnte man sich einen Staffellauf vorstellen, indem der Stab immer weiter gegeben wird. Ein Token kann durch Übertragungsfehler oder den Ausfall von Stationen verloren gehen. Da auch hier der Datenverlust sehr hoch ist, wird diese Topologie nicht mehr verwendet [LIN13].

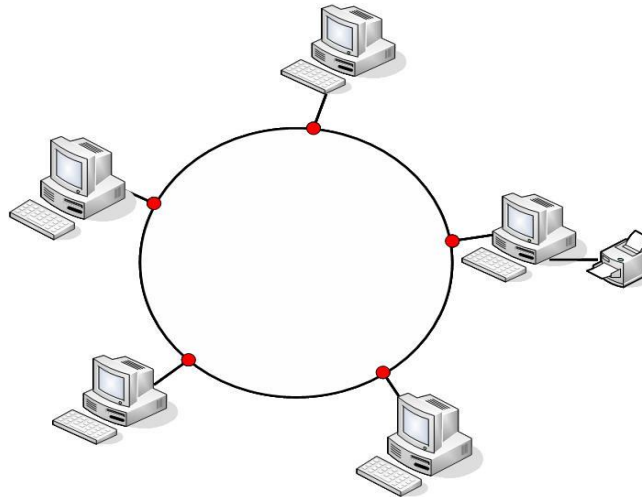


Abbildung 5 Schema einer Ringtopologie [BOD06]

2.1.2.1.3 Sterntopologie

Bei dieser Anordnung gibt es ein zentrales Element, den Hub, von dem sternförmig die Leitungen zu den einzelnen Geräten führen. Ein Hub ermöglicht die Kommunikation zwischen zwei Rechnern in einem Netzwerk. Er ist nicht im Stande, die Quelle oder das gewünschte Ziel der erhaltenen Informationen zu identifizieren und sendet alle Informationen an alle angeschlossenen Geräte (auch an die Quelle). Er kann nicht gleichzeitig senden und empfangen [MIC14]. Jedes Endgerät verfügt über eine eigene Leitung, wodurch es von Störungen auf anderen Leitungen unangetastet bleibt. Diese Topologie findet bei Fast- und Giga-Ethernet Verwendung [LIN13].

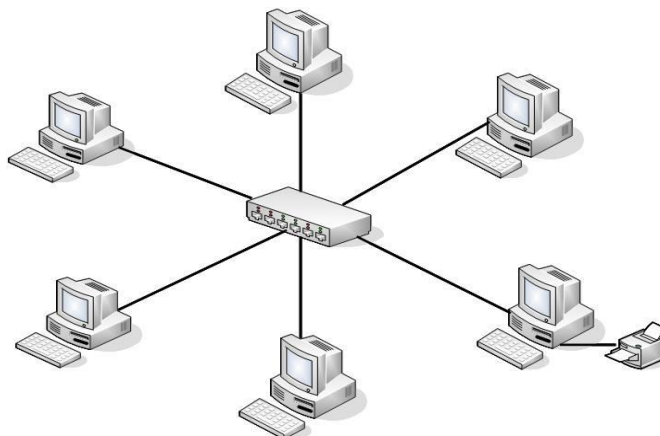


Abbildung 6 Schema einer Sterntopologie [BOD06]

2.1.2.2 Lokale Netze (LANs)

Lokale Netze beschränken sich auf ein Gebäude oder ein Gelände und besitzen eine Reichweite von wenigen Kilometern. Sie finden oft für private Zwecke Anwendung, wie das Verbinden von PCs und Workstations (z.B. Drucker). In den meisten Fällen basiert die Übertragung auf einem Kabel, an das alle anderen Geräte angeschlossen sind. Die Übertragungsgeschwindigkeiten herkömmlicher LANs belaufen sich zwischen 10 – 100 Mbit/s – moderne arbeiten mit 10Gbit/s [TAN03].

2.1.2.2.1 Ethernet

Die Urform des heutigen Ethernets geht auf Bob Metcalfe und David Boggs zurück, die es auf Grundlage der Forschungen von Norman Abramson entwickelten. Norman Abramson versuchte den Umstand, dass es Anfang der 70er Jahre auf Hawaii kein funktionierendes Telefonnetz gab, zu egalisieren, indem er über den Kurzstreckenfunk alle Inseln mit dem Hauptrechner auf Honolulu verband [TAN03].

„Benannt ist es nach dem Lichtäther, von dem einmal angenommen wurde, dass sich in ihm elektromagnetische Strahlung fortpflanzt“ [TAN03 Seite 85]. Als Übertragungsmedium fungiert ein in den Anfangsjahren ein Koaxialkabel (Äther), an das maximal 256 Rechner angeschlossen werden konnten (Bustopologie) und eine Übertragungsrate von 2,94 Mbit/s hatte [TAN03].

Heutzutage werden Twisted Pair-Kabel verwendet. Der Vorteil, der sich hieraus ergibt, besteht darin, dass je ein Adernpaar zum Senden und Empfangen benutzt wird, was den Nutzen hat, dass ein Gerät gleichzeitig senden und empfangen kann (Vollduplex) [LIN13]. Zudem hat man sich von der Bustopologie hin zur Sterntopologie mit Hub entwickelt. In den frühen 2000er Jahren wurde der Hub durch einen Switch (teilt ein Ethernet in mehrere kollisionsfreie Segmente auf) ersetzt [KUR08].

1995 wurde das Fast-Ethernet mit einer Geschwindigkeit von 100 Mbit/s normiert (durch die Arbeitsgruppe der IEEE). Die Daten werden bei dieser Variante dichter übertragen, wodurch die Laufzeit einer Informationseinheit deutlich gesenkt wird [LIN13].

Gigabit-Ethernet bietet Geschwindigkeiten bis zu 1000 Mbit/s, was aber von einem Rechner selten stetig ausgenutzt wird. Anwendung findet es bei Backbones (Hauptnetzverbindungen in großen LANs). Als Übertragungsmedium werden häufig Glasfaserkabel verwendet [LIN13].

2.1.2.2.2 WLAN

Die WLAN-Technologie zählt heute sicherlich zu den bedeutendsten Netzwerkszugangstechnologien, da sie fast überall den Zutritt zum Internet ermöglicht. Soziale Netzwerke und Microbloggingdienste haben einen großen Anteil ihrer Popularität der drahtlosen Technologie zu verdanken, da man nicht direkt vor dem Rechner sitzen muss, sondern aktuelle Fotos, Statements oder persönliche Eindrücke von dem Ort, an dem man sich gerade befindet, posten kann und auch über Aktivitäten von befreundeten Personen sofort auf dem Laufenden gehalten wird.

Um diese Technik nutzen zu können, benötigt der Computer ein Funkmodem um mit den anderen Systemen kommunizieren zu können. WLANs haben sich in den letzten Jahren immer mehr in Büros, Flughäfen oder anderen öffentlichen Behörden etabliert als Ethernets, da die Installation nicht so aufwändig ist [TAN03].

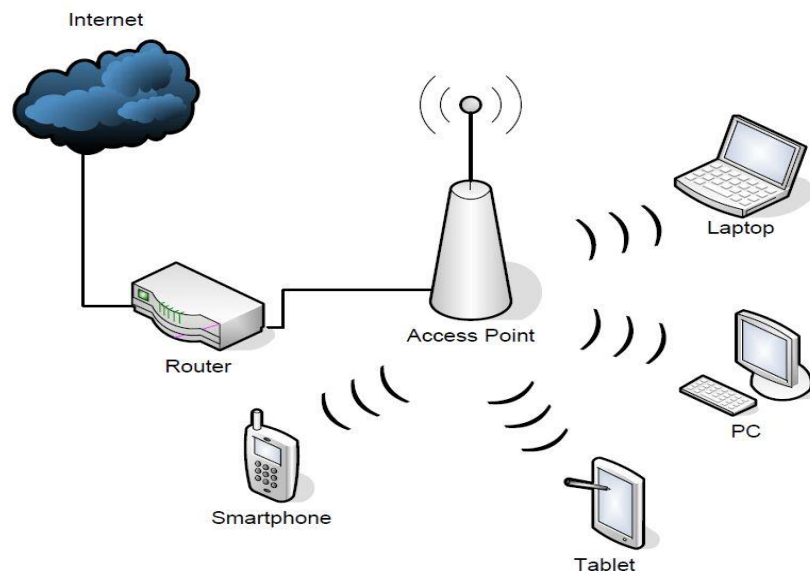


Abbildung 7 Schema WLAN [WLA14]

Darüber hinaus gibt es noch den Kurzstreckenfunk Bluetooth (benannt nach dem Wikingerkönig Harald Blaatand (Bluetooth) II, der Dänemark und Norwegen vereinte), dessen Entwicklung auf das norwegische Unternehmen L.M. Ericsson zurückgeht [TAN03]. Dieser funktioniert nach dem Master-Slave-Prinzip: Der Computer ist in der Regel der Master, der mit allen peripheren Geräten (z.B. Tastatur, Maus, Drucker, Kopfhörer u.a.), den Slaves kommuniziert. Bluetooth-Netzwerke sind Ad-hoc Netzwerke, wofür man keinen Access Point benötigt. Alle Geräte werden in einem Piconetz (max. 8 aktive Geräte) organisiert (1 Masterknoten, max. 7 Slavesknoten), das vom Master ausgeht und eine maximale Reichweite im Umkreis von 10m hat [KUR08].

Der große Vorteil dieser Technologie ist, dass am PC-Arbeitsplatz ein Kabelsalat verhindert wird, da diese nun überflüssig geworden sind. Des Weiteren sind sie sehr energieschonend.

Im Normalfall reicht es aus, an allen beteiligten Geräten das Bluetooth einzuschalten und sie erkennen sich untereinander, ohne vorangegangene Treiberinstallation [TAN03].

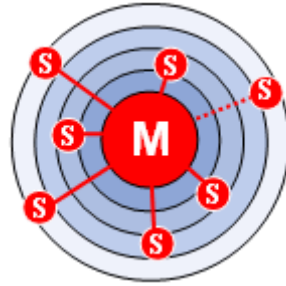


Abbildung 8 Kommunikationsprinzip Bluetooth [KIO14]

2.1.2.3 Weitverkehrsnetze

In dieser wissenschaftlichen Arbeit wird in diesem Kapitel nur konkret das Internet als Weitverkehrsnetz dokumentiert. Das Internet kann nicht als eigenständiges Netz, sondern als Ansammlung vieler verschiedener Netze betrachtet werden. Die nachfolgenden Technologien bilden die Grundlage des Internets:

- **MAN** (Metropolitan Area Network): Stadtnetz z.B. Kabelfernsehtznetz, Verbindung vieler LANs
- **WAN** (Wide Area Network): Fernnetz, Ausdehnung auf einen großen geographischen Bereich (Land/Kontinent)
- **GAN** (Global Area Network): keine geografische Begrenzung, Verbindung vieler WANs

Das Internet ist ein Computernetzwerk, das Millionen von Rechnern weltweit miteinander verbindet. Es ist ein paketvermitteltes Netz (übermittelte Daten werden in Pakete zerlegt und beim Zielsystem wieder zusammengesetzt). Mittlerweile zählen aber nicht nur Rechner zu den Geräten, die mit dem Internet verbunden sind, sondern auch Smartphones, Webcams, TV-Geräte oder Fahrzeuge. Diese Geräte werden Hosts genannt und die Anzahl der Geräte, die mit dem Internet verbunden sind wachsen kontinuierlich weiter. Die Endgeräte sind über die oben genannten Übertragungsmedien miteinander verbunden [KUR08].

Möchten Endgeräte auf das Internet zugreifen, benötigen sie einen Internetdienstanbieter (*Internet Service Provider*), der ihnen den Zugang ermöglicht. Dieser Netzzugang kann über 56 Kbit/s-Einwahlmodem (veraltet), Breitbandzugänge (z.B. DSL) mit Routern als Hardwarekomponente, lokale Hochgeschwindigkeitsnetzwerke oder über drahtlose Zugänge erfolgen [KUR08].

Alle Endgeräte nutzen zum Senden und Verschicken von Daten Protokolle. Die wichtigsten sind dabei TCP (Datenübertragungsprotokoll) und IP (legt Format der Pakete fest, die

zwischen Router und Endgerät versendet und empfangen werden) [KUR08]. Neben den Protokollen benötigt man einen Switch beziehungsweise einen Router um die Pakete an ihren Zielort weiterzuleiten. Ein Router ist eine Vermittlungseinheit, die Daten, die auf einer Eingangsleitung eingehen auf eine Ausgangsleitung weiterleiten [TAN03]. Switches ähneln Hubs in ihrer Funktionsweise. Sie können jedoch das vorgesehene Ziel identifizieren an welches die Informationen gesendet werden sollen und können gleichzeitig senden und empfangen. Sie sind daher schneller als ein Hub, jedoch teurer in ihrer Anschaffung [WIN14]

In den nachfolgenden Kapiteln werden einige Kommunikationsanwendungen, die über das Internet möglich sind näher erläutert. Ferner zählt auch das Websurfen, Internetradio und –fernsehen sowie Videostreaming zu beliebten Internetanwendungen. Sie tragen den Namen verteilte Anwendungen, weil sie mehrere Endgeräte zur Ausführung benötigen [KUR08].

2.2 Computervermittelte Kommunikationsformen

2.2.1 Asynchrone computervermittelte Kommunikation

Asynchrone computervermittelte Kommunikation bedeutet, dass keine zeitgleiche Anwesenheit von Sender und Empfänger notwendig ist. In den nachfolgenden Kapiteln werden die wichtigsten asynchronen Kommunikationsformen beschrieben.

2.2.1.1 E-Mail

1971 verschickte der Programmierer Raymond Samuel Tomlinson die erste E-Mail zwischen zwei Rechnern, die über das Arpanet miteinander verbunden waren an seine eigene E-Mail Adresse und definierte zudem die Bedeutung des @-Zeichens. Eric Allman programmierte 1981 die Software Sendmail, die es dem Benutzer erlaubte, Nachrichten mit einem Mailprogramm gleichzeitig in verschiedene Netze zu verschicken [BEC06].

Als der E-Mail-Dienst entwickelt wurde, war er eine schnellere Alternative als die Briefpost („Snail Mail“ = Schneckenpost). Des Weiteren ist er kostengünstiger und weniger aufdringlich wie das Telefon und gegenüber dem Telefax liegen die Dokumente elektronisch vor und können direkt weiterverarbeitet werden [DÖR06].

Der E-Mail Dienst basiert auf den 3 Protokollen SMTP, POP3 und IMAP:

- **SMTP:**
Ein Protokoll, das dem Austausch von E-Mails ins Computernetz dient. Besonders beim Versenden und Weiterleiten der E-Mail an den E-Mail Server ist es von großer Bedeutung [LIN13].
- **POP3:**
Ein Protokoll, das mit Hilfe eines E-Mail Clients (z.B.Thunderbird oder Outlook) das

Abholen und lokale Speichern der E-Mails von einem E-Mail Server ermöglicht. Dies ist jedoch erst möglich, nachdem der Benutzer sich am Server des E-Mail-Providers angemeldet und authentifiziert hat. In der Regel stehen die E-Mails nach dem Abholvorgang nur noch lokal auf dem E-Mail-Client zur Verfügung. Sollte jedoch ein zweites Gerät hinzukommen, dass ebenfalls die E-Mails vom Server des Providers abholen will, kann der E-Mail Client nur diejenigen abholen, die noch nicht vom ersten Gerät abgeholt wurden [LAN13].

- **IMAP**

Protokoll, welches den Zugriff mit mehreren Endgeräten auf den Mailbestand zulässt. IMAP ist die Erweiterung zu POP. Alle Mails bleiben auf dem Server des E-Mail-Providers und müssen auch nicht runtergeladen werden, wenn man die Mails lesen möchte. Die Einteilung nach bestimmten Kriterien wie „Gelesen“, „Beantwortet“ oder „Entwurf“ erleichtern dabei den Überblick. Nachteil ist dabei, dass der E-Mail-Provider ein großes Speichervorkommen bereitstellen muss, da jeglicher E-Mail Verkehr auf dem Server gespeichert wird [LAN13].

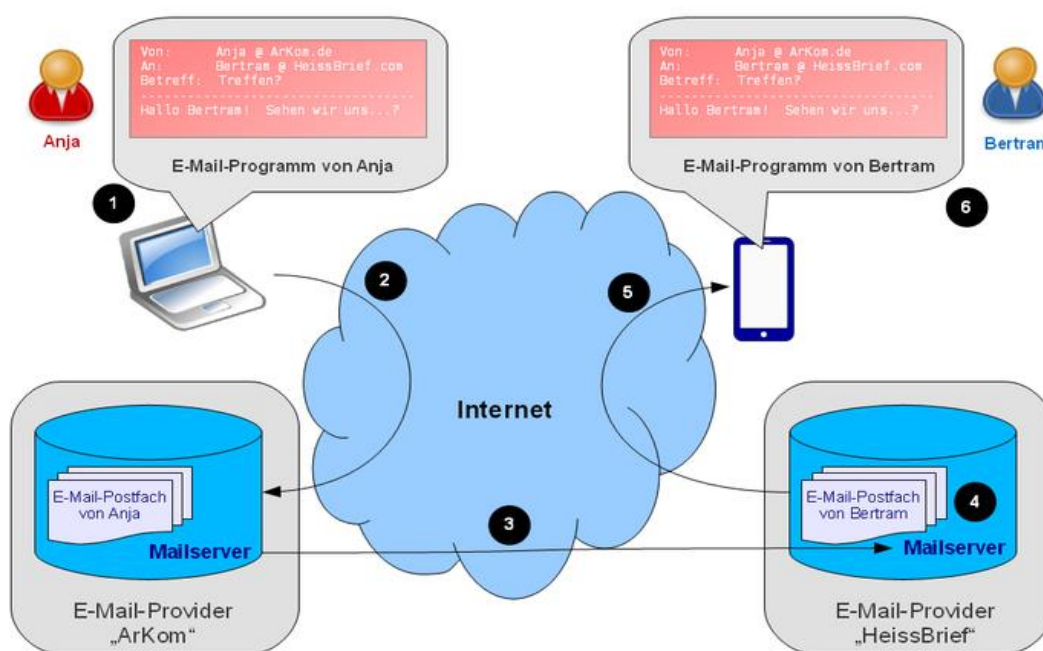


Abbildung 9 Schema: prinzipieller Ablauf einer E-Mail Zustellung [WIK143]

Eine E-Mail besteht in jedem Fall aus einem Header sowie dem Body und kann ebenfalls noch Anhänge (Attachments) enthalten. Im Header finden sich Informationen über den Übertragungsweg der E-Mail, E-Mail-Adressen des Absender und Empfängers, Betreffzeile und das Sendedatum mit Uhrzeit. Der Body enthält die zu übermittelnde Nachricht und eventuell Signaturen, die Aufschluss über den Empfänger geben können, ähnlich einer Visitenkarte. Als Attachment versteht man eine Datei, die zusätzlich an die versendete E-Mail angehängt wurde. Dafür eignen sich Bilder, Dokumente, aber auch komprimierte Daten [BEC06].

Die E-Mail-Adresse besteht aus mehreren Bestandteilen. Bei *mkonietz@hs-mittweida.de* ist *hs-mittweida* die Sub-Domain und liefert Informationen über das organisationsinterne Netz, in dem sich der Rechner befindet. Die Top-Level-Domain *de* verweist auf das Land, in dem der Rechner vernetzt ist. Der Benutzername *mkonietz* dient der individuellen Beschreibung der E-Mail-Adresse [STE06].

Laut einer Onlinestudie von 2010, zeigt sich an der Tendenz, dass die E-Mail in den kommenden Jahren von den sozialen Netzwerken als meistgenutztes Online-Kommunikationsmittel abgelöst wird. Schon jetzt verbringen die Online-User in den Ländern, in denen die Befragung stattfand mehr Zeit mit dem Kommunizieren über soziale Netzwerke als mit dem Schreiben und Lesen von E-Mails [TIL10].

2.2.1.2 Mailinglisten

Mailinglisten stellen im Gegensatz zur Rundmail ein zentral organisiertes Forum dar, das nicht auf einen Verteiler zurückgreift, sondern auf ein softwaregestütztes Programm welches die Empfängeradressen verwaltet. Die häufigste Umsetzungsform ist dabei die E-Mail [BEC06]. Ein Beispiel anhand der E-Mail würde bedeuten, dass eine Liste aller E-Mail-Adressen der Mitglieder existiert, die dann über eine Sammeladresse (Mailinglistenadresse) erreichbar sind. Sobald eine Person mit seiner E-Mail Adresse in der Mailingliste eingeschrieben ist, erhält sie sämtliche Nachrichten, die über die Sammeladresse versendet werden. Diese Botschaften heißen Postings oder Posts um eine Abgrenzung zur E-Mail zu schaffen [DÖR06].

2.2.1.3 Newsgroups

Eine weitere Gruppe asynchroner Kommunikationsformen sind die Newsgroups, die ebenso wie Mailinglisten, darauf abzielen, elektronische Inhalte vielen Nutzern gleichzeitig in virtuellen Internetforen zur Verfügung zu stellen, die dann die Möglichkeit haben, die Beiträge zu kommentieren [FRA12]. Usenet war 1979 dabei das erste Forum dieser Art und war im Gegensatz zum Arpanet eine gebührenfreie Alternative, die ihr eigenes Protokoll (UUCP) nutzte. Zu Beginn verband es nur zwei Rechner miteinander, den der Duke University und den der University of North Carolina, und besaß 15 Newsgroups [BEC06]. Seit den 2000er Jahren werden sie aber immer mehr von Webforen ersetzt

Um Zugriff auf Newsgroups zu erhalten, benötigt man einen bestimmten Newsreader (Client) oder eine Webschnittstelle (z.B. Google Groups). Diese Newsreader-Clients sind heute als Teilprogramme in den meisten E-Mail-Programmen integriert (z.B. Mozilla Thunderbird) oder als Add-on in Webbrowsern enthalten. Der Newsreader stellt dann die Verbindung zum Newsserver her und der Nutzer ist nun im Stande, die dort verfügbaren Newsgroups zu nutzen. Das Senden und Empfangen der Posts wird über die Protokolle UUCP oder über das IP basierte NNTP realisiert [DÖR06].

Man unterscheidet hierbei zwei unterschiedliche Arten. Die unmoderierten Gruppen, in denen jeder Beitrag sofort für alle Teilnehmer sichtbar ist, sowie die moderierten Gruppen, in denen jeder Beitrag zuerst bei einem beziehungsweise mehreren Moderatoren eintrifft und erst nach deren Freigabe für alle sichtbar wird [BEC06].

2.2.1.4 Internetforen

Das oben erwähnte Usenet sowie die Mailinglisten können als Vorgänger der heutigen modernen Internetforen angesehen werden, in denen Gedanken und Erfahrungen in einen virtuellen Bereich ausgetauscht werden können [DÖR06].

In der Regel gibt es in jedem Forum ein Hauptthema (z.B. Fernsehserien), welches sich dann in die jeweiligen Unterthemen (z.B. Soaps, Krimis, Comedy u.a.) spezialisiert. Die verfassten Beiträge werden, wie in der Mailinglist, als Postings bezeichnet. Der sich darauf aufbauende Diskussionsfaden mit anderen Teilnehmern wird Thread oder Topic genannt [MÜN07]. Wie auch in den Newsgroups kann es vorkommen, dass manche Themenbereiche durch einen Moderator geleitet werden.

Die Anmeldung und Nutzung vieler Internetforen ist kostenlos. Die sogenannte Supportforen, in denen User nach Ratschlägen und Tipps zum „Selberhelfen“ suchen können, oder diverse Game-Support-Foren haben einen kommerziellen Hintergrund.

Ein Begriff, der häufig im Zusammenhang mit Internetforen fällt, ist Bulletin Board. Darunter kann man sich ein schwarzes Brett vorstellen in Form eines offenen Diskussionsforums, auf dem Informationen abgelegt werden können und von mehreren Benutzern gelesen werden können [SPR].

2.2.1.5 Blog

Ein Weblog ist meistens ein frei zugängliches Online-Tagebuch, das von einer Person (Blogger) oder einer Organisation geführt wird. Der Blogger führt dabei eine endlose chronologische Liste in der er Sachverhalte protokolliert oder Gedanken niederschreibt, die umgekehrt chronologisch dargestellt werden. In vielen Blogs ist es möglich, dass weitere Personen Kommentare über das Niedergeschriebene verfassen können [FRA12].

Als erstes Blog gilt die Website des WWW-Entwicklers Tim Berner Lee, die 1990 online ging und auf der er fortlaufend den Entwicklungsstand der Internettechnologie protokollierte. 1997 wurde der Begriff Weblog vom US-Amerikaner Jorn Barger erstmals in seinem Netztagebuch verwendet. Pitas stellte 1999 in den USA die freie Weblogsoftware dar [NET]. Mittlerweile bieten diverse soziale Netzwerke, wie MySpace, ihren Mitgliedern die Blogfunktionalität an.

Ein Blog kann aus mehreren Elementen bestehen [EDU05]:

- **Beiträge:** Hauptbestandteil des Blog, chronologisch umgekehrt aufgelistet
- **Thread:** Gesamtheit aller Beiträge
- **Permalinks:** permanente Webadresse die auf ein einzelnes Weblog verweist, auch wenn dieser bereits archiviert wurde.
- **Kommentar:** eigene Meinung/Feedback zum ausgewählten Beitrag
- **Trashback/ Pingback:** gibt Bloggern die Möglichkeit festzustellen ob in einem anderen Beitrag Bezug auf den Eigenen genommen wurde.
- **RSS-Feed:** informiert Leser über neu erschienene Beiträge im Blog
- **Blogroll:** Linksammlung zu anderen Weblogs
- **Asides:** kleine Einträge, die nur aus wenigen Wörtern oder Zeilen bestehen
- **Sphäre:** Gesamtheit aller Weblogs/Blogger (Blogosphäre)

Blogs stellen heute ein nützliches und schnelles Kommunikationsmittel dar. So nutzte der chinesische Künstler Ai Weiwei aufgrund der großen Zensur, die in seinem Land vorherrscht, Blogs als Kommunikationsmittel nach außen, um auf die Missstände Chinas aufmerksam zu machen. Viele Prominente nutzen Weblogs für PR-Zwecke, um zum Beispiel ihre Bekanntheit zu steigern. Tendenziell steigt die Zahl von Weblognutzern jährlich.

2.2.2 Synchrone computervermittelte Kommunikation

Im Gegensatz zur asynchronen Kommunikationsform müssen bei der synchronen die Kommunikationspartner zur selben Zeit aktiv sein um die ausgetauschten Botschaften relativ zeitgleich zu beantworten [DÖR06].

Der Vorteil der sich hieraus erschließt besteht darin, dass die übertragenen Botschaften zeitlich unmittelbar beim Kommunikationspartner eintreffen. Hier wird aber auch der Nachteil sichtbar: während bei der E-Mail „Sender und Empfänger den Zeitpunkt ihrer kommunikativen Aktivität selbst bestimmen können“, [DÖR06 Seite 53] müssen beide Partner bei der synchronen Kommunikationsart aktiv sein. Diese Kommunikation muss aber nicht nur zwischen zwei Personen stattfinden, sondern kann auch mehrere Kommunikationspartner mit einbeziehen [DÖR06]

2.2.2.1 Internet-Telefonie

Die erste Realisierung digitale Sprache zwischen zwei Rechnern, die über das Arpanet miteinander verbunden waren, umzusetzen, erfolgte 1973. Auch Mitte der 90er Jahre war die Technik noch nicht ausgereift, weil die Datenleitungen über den schmalbandigen Modem oder ISDN-Anschluss kein wirkliches Telefonieren zuließen. Der Durchbruch kam mit den breitbandigen DSL-Leitungen. Internet-Telefonie ist ebenfalls unter den Namen IP-Telefonie oder Voice over IP (weil das Telefonieren mittels Internet-Protokoll (IP) realisiert wird) bekannt [NÖL05].

Der entscheidende Vorteil dieser Technologie besteht in der Kosteneinsparung, da ein Netz für alle Dienste verwendet wird. Nachteilig wirkt sich die paketvermittelte Übertragung aus, da Datenpakete verloren gehen können und es so zu Sprachaussetzern, Echos, Rauschen oder verzerrten Stimmen kommen kann. Diese Probleme konnten aber in den letzten Jahren deutlich minimiert werden [NÖL05].

Funktionell muss zuerst ein analoges Sprachsignal mit einem Mikrofon aufgenommen werden. Mithilfe eines Analog/Digital-Wandlers werden die analogen Signale in ein digitales umgewandelt und in das entsprechende Audioformat kodiert. Die Sprachqualität ist dabei vom gewählten Komprimierungsverfahren abhängig. Nachdem die Daten in das entsprechende Format kodiert wurden, werden sie über das Netzwerk zum Kommunikationspartner gesendet. Diese Übertragung geschieht paketvermittelt. Beim Empfänger werden diese Pakete zunächst in einem Puffer zwischengespeichert bevor sie den Digital/Analog-Wandler durchlaufen und am Lautsprecher in die menschliche Sprache zurückgewandelt werden [TIK].

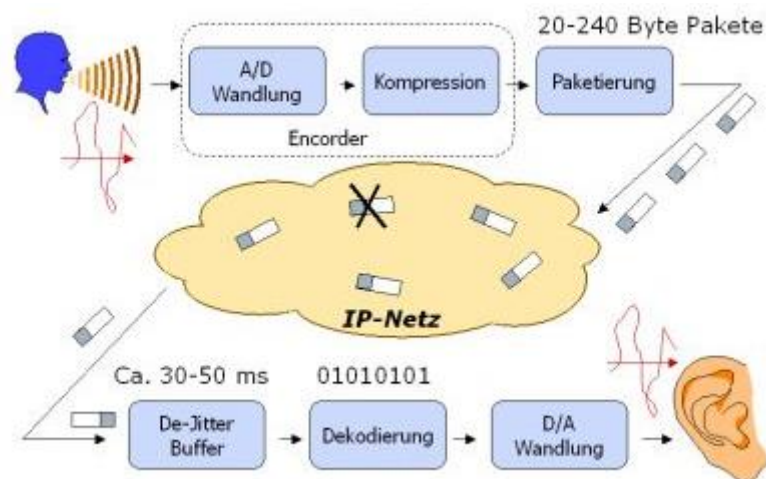


Abbildung 10 Funktionsprinzip Internet Telefonie [MOO14]

Möchte man eine Verbindung zu einem gewöhnlichen Festnetzanschluss herstellen nutzt man ein Gateway. Diese Vermittlungsrechner sind zum einen mit dem Internet und zum anderen mit dem normalen Telefonnetz verbunden und können somit Telefonate in beide Richtungen weiterleiten [TIK].

Zum Marktführer in Sachen Voice over IP hat sich in den letzten Jahren die 2003 erschienene Skype-Software entwickelt, mit der Sprach- und Videoanrufe zu anderen Skypebenutzern kostenlos sind. Das Unternehmen Google versucht aber mit seinem Dienst Google Voice, das bei Google Mail integriert ist, diese Vormachtstellung zu beenden

2.2.2.2 Chats

Das erste Chat(IRC)-Programm wurde im August 1988 vom finnischen Studenten Jarkko Oikarinen entwickelt. Aktuell gibt es mehrere tausend verschiedene Kanäle, die es Personen ermöglichen weltweit über die verschiedensten Themen miteinander zu chatten. Die meisten Unterhaltungen finden im offenen Bereich des Chatrooms statt. Es ist jedoch möglich andere User in kleinere Chatrooms einzuladen, um dort Dialoge zu führen. Die Kommunikation geschieht in der Regel pseudonym, da die meisten User einen Nickname verwenden. Das chatten kann in organisatorisch, thematisch und institutionell geregelten Channels stattfinden [BEC06]:

IRC-Chats:

Diese Channels bilden den größten nicht-kommerziellen Dienst. Die Kanäle werden durch ein vorangestelltes # symbolisiert, in der es keine hierarchisch-thematische Ordnung gibt [BEC06]. Grundlage ist dabei eine Client-Server-Architektur, die die Installation eines IRC-Client voraussetzt. Das erste nutzbare IRC-Netzwerk war EFnet, welches Mitte der 90er Jahre entstand und mit der zunehmenden Zahl an Usern und Channels Defizite in der Datenübertragung verkraften musste [DÖR06].

Dem Benutzer steht es frei, ob er im öffentlichen Bereich des Chatrooms kommuniziert oder kleinere Chatrooms anlegt und weitere User hierin einlädt. Wie bereits oben erwähnt, muss sich der anmeldende User einen Nickname aus maximal 9 Zeichen wählen und kann dann über das Kommando „join“ an einem Chat teilnehmen. Diese Channels stehen dem User 24 Stunden am Tag zur Verfügung [DÖR06].

Viele verschiedene Interessen werden in den Channels bedient. So finden Anhänger gleicher Fankulturen, Szene- oder Subkulturen sowie Personen mit lokalen Bezügen gleichgesinnte Kommunikationspartner. In diesen offenen Channels ist es möglich unpassende Beiträge zu unterbinden [BEC06].

Dem gegenüber stehen die geschlossenen Chats. Diese werden moderiert und besitzen vorgegebene Themen. User, die sich nicht an die Vorgaben halten, können im schlimmsten Fall aus dem Chatroom ausgeschlossen werden. Moderatoren erkennt man am @-Zeichen vor ihrem Nickname [BEC06].

Web-Chats:

Diese Art Channel etablierte sich Mitte der 90er Jahre, als sich das Interesse am World Wide Web steigerte. Für diese Kanäle benötigt man keine zusätzliche Client-Software, da man sie direkt über einen Browser nutzen kann, indem man die entsprechende URL eingibt [DÖR06].

Neben den normalen Nachrichten in Textform halten auch immer mehr visuelle Gestaltungsformen Einzug in die Kommunikationen. Unter anderem äußert sich dies mit Avataren (bildliche Repräsentationen der Chatter). Diese spezielle Form trägt den Namen

Grafik-Chat und kann als 2D- oder 3D-Welt existieren. In der 3D-Welt ist es für die vorgegebenen beziehungsweise selbst erstellten Avatare möglich, sich frei in der 3D-Welt zu bewegen und sich mit anderen Avataren zu unterhalten. Da sie sehr hohe Bandbreiten brauchen und zum Teil sehr unkomfortabel sind konnten sie sich noch nicht durchsetzen [DÖR06].

Großer Beliebtheit erfreuen sich einmalige Chats mit Prominenten oder Experten. Diese Events werden moderiert und haben primär den Zweck der wissenschaftlichen Analyse, da die Protokolle archiviert werden [DÖR06].

Daraus entstanden dann auch die Talkdienste, in denen konkret ein Gesprächspartner angeschrieben werden muss. Dem angeschriebenen Gesprächspartner steht es frei ob er die Konversation annimmt oder nicht. Der bekannteste Talkdienst ist sicherlich ICQ. Dieses Programm sucht, nachdem man online ist, sofort nach Bekannten, die ebenfalls online sind. Kommerzielle Varianten wären die Instant Messenger verschiedener Online Dienste [BEC06].

2.2.2.3 MUDs und MOOS

1979 wurde das erste MUD (Multi User Dungeon/Domain beziehungsweise Dimensions) von zwei englischen Studenten an der Essex University erstellt. „Technisch betrachtet sind MUDs softwaregesteuerte Datenbanken, deren Inhalt aus schriftsprachlichen Beschreibungen der Figuren, Räume und Objekte besteht sowie über Zu- und Ausgänge zu den beschriebenen Umgebungen“ [BEC06 Seite 135] verfügt – der Nutzer begibt sich in eine textbasierte virtuelle Realität, um mit anderen Usern auf Grundlage eines Rollenspiels zu kommunizieren.

Dabei unterscheidet man zwischen Abenteuer-MUD, wo der User sich in eine Fantasiewelt begibt und dort verschiedene Aufgaben bewältigen muss (Rätsel lösen, Feinde bekämpfen, magische Kräfte erwerben u.a.) und dem sozialen MUD, bei dem das Augenmerk eher auf der zwischenmenschlichen Interaktion zwischen zwei Usern liegt. In beiden Varianten gestalten die User einen Avatar mit einem Nickname, der stellvertretend den User in der Fantasiewelt repräsentiert [DÖR06].

MOO (MUD Object-Oriented) sind spezielle MUDs, die es dem User ermöglichen ausdifferenzierte Formen der sozialen Interaktion zu ermöglichen sowie das Erzeugen von virtuellen Räumen oder Objekten zu erleichtern [DÖR06].

Im Gegensatz zu E-Mail und WWW ist die Nutzerzahl sehr gering, da diese synchrone Kommunikationsform nur eine kleine Fankultur anspricht.

2.2.3 Soziale Netzwerke und Personal Publishing

Soziale Netzwerke bekommen ein eigenes Unterkapitel, da man bei ihnen nicht konkret nach synchroner oder asynchroner Kommunikation differenzieren kann, da viele der oben beschriebenen Dienste und Anwendungen in den sozialen Plattformen integriert sind. So ist es bei Facebook möglich Nachrichten in Form einer E-Mail an Benutzer zu senden, Videoanrufe zu tätigen oder mit Freunden und Bekannten zu chatten, die gleichzeitig online sind. Ebenso das Publizieren multimedialer Inhalte steht im Vordergrund solcher Plattformen.

Das Wort Netzwerk im Namen weist dabei aber weniger auf ein technisches Netzwerk hin, wie es in Kapitel 2.1.2 beschrieben ist, als vielmehr ein Netzwerk von Beziehungen zwischen Benutzern (z.B. Freundschaften).

Wie bei den sozialen Netzwerken lässt sich diese Differenzierung auch nicht auf Dienste des Personal Publishing beziehen. Beispiele hierfür sind Microbloggingdienste (z.B. Twitter) sowie Podcasts oder Videocasts, weil sie ebenso asynchrone wie synchrone Kommunikationsformen vereinen.

3 Der Microbloggingdienst Twitter

3.1 Microblogging

Twitter kann nicht als soziales Netzwerk als solches angesehen werden, sondern wird in der Kategorie Microbloggingdienst klassifiziert. Während normale Blogs keine Zeichenbegrenzung haben, sich aber aus Gründen der Leserkonzentration auf maximal 500 Wörter konzentrieren, gibt es bei Twitter die Einschränkung auf maximal 140 Zeichen, was aus dem Namen Microblogging, also „verkleinertes“ Bloggen, hervorgeht. [SIM10].

Wenn ein Nutzer eine Nachricht (micro)bloggen möchte, kann er dies mit einem Software-Client (z.B. TweetDeck), einem Instant Messenger, per SMS oder direkt zu Hause vom Rechner machen. Die Nachrichtenlänge von maximal 140 Zeichen speziell bei Twitter ergibt sich aus den vielen SMS-basierten Empfangsfunktionen: Eine SMS darf maximal 160 Zeichen haben, zieht man eventuell eingegeben Befehle und den Username ab verbleiben noch ca.140 Zeichen für die eigentliche Mitteilung [SIM10]. Diese Begrenzung wurde eingeführt, damit SMS-Nachrichten auch über das Internet verbreitet werden können, da in den USA, dem Entwicklungsland von Twitter, der Empfänger einer SMS ebenfalls eine Gebühr zahlen muss. Twitter bietet dem Sender die Möglichkeit einen Tweet an alle beteiligten Personen zu schreiben, ohne beim Empfänger Kosten zu verursachen. Der Verfasser des Tweets wird somit zu einer kurzen und präzisen Aussage zum beschriebenen Thema gezwungen ohne tiefgreifender in der Thematik zu werden. In der Regel handelt es sich bei den Kurznachrichten um reine Textbotschaften, die keine Bilder, Videos oder Musikdateien enthalten, was jedoch nicht ausgeschlossen ist [SIM10].

Neben dem Marktführer Twitter gibt es noch Alternativen, die zum Teil durch andere Funktionalitäten erweitert sind. So sind die Microbloggingdienste Co-op und Yammer darauf ausgelegt die Kommunikation in Unternehmen zu erleichtern („Business-Microblogging“) zum Beispiel für eine schnelle Projektkoordination beziehungsweise Wissensaustausch unter Mitarbeitern. Obayoo von Apple dient ebenfalls der Unternehmenskommunikation und gestattet die Archivierung von Nachrichten. Neben den bereits genannten wären noch Sharetronix, Snipia, Socialcast und StatusNet erwähnenswert, die in erster Linie für den Unternehmensgebrauch erdacht sind, aber auch im privaten Gebrauch Anwendung finden [COM11].

3.2 Twitter

Da sich das Programm TwitterSearcher, wie es der Name schon sagt, nur auf die forensische Datensuche in Twitter beschränkt, wird ausschließlich in dieser wissenschaftlichen Arbeit dieser Microbloggingdienst näher erläutert, welcher 2006 von den Webprogrammierern Evan Williams, Biz Stone und Jack Dorsey entwickelt wurde [SIM10]. Twitter bietet Menschen die Möglichkeit, Meinungen in Form von Tweets auszudrücken und

mit Freunden zu kommunizieren. Ein entscheidender Vorteil, der sich in den letzten Jahren herausgestellt hat, besteht darin, dass weltweite Ereignisse, zum Teil bevor Bekanntwerden in Nachrichtensendungen, schon bei Twitter gepostet wurden. Aus diesem Grund haben Nachrichtenagenturen großes Interesse, die Kommunikation bei Twitter zu überwachen, was diesen Trend in den nächsten Jahren weiter vorantreiben wird. Zudem ist Twitter für viele Unternehmen ein wichtiges Instrument für Marktforschungen geworden.

Das forensische Analysewerkzeug untersucht ausschließlich Nachrichten bei Twitter, weil Twitter entgegen anderen Plattformen eine frei zugänglich API (Programmierschnittstelle) zur Verfügung stellt, mit der es möglich ist auf Inhalte des Microbloggingdienstes zuzugreifen. Diese API ermöglicht Programmentwicklern und Webprogrammierern Zugriff auf Twitter oder einzelne Tweets, um diese in eigenen Softwareprogrammen oder Websites einzubinden. Selbstverständlich besitzt auch das soziale Netzwerk Facebook eine API („Graph-API“), diese ist aber auch in ihren Zugriffsmöglichkeiten sehr eingeschränkt und stellt nur frei zugängliche Informationen zur Verfügung. Verglichen mit der Twitter-API bedarf es bei der Facebook-API auch einer längeren Einarbeitungszeit durch den Entwickler, da diese viel komplexer ist. Offene APIs bieten den Betreibern im Zeitalter von Web 2.0 („Webinhalte werden von den Internetbenutzern selbst bereitgestellt und miteinander geteilt, um einen sozialen Kontext zu erstellen“ [SIM10 Seite 16] z.B. das Posten und Verwalten von Bildern, der Austausch von Videos oder Microblogging) neue Marketingmöglichkeiten, da durch das Einbinden in Drittprogramme die Popularität des eigentlichen Programmes steigt.

Sicherlich würde Facebook, das beliebteste soziale Netzwerk in Deutschland, oder auch diverse lokale Netzwerke bei der Ermittlern mehr Interesse wecken, da es hier eine größere Nutzergemeinde (27 Millionen, Stand Januar 2014 [STA14]) als bei Twitter gibt, beziehungsweise in den lokalen Netzwerken sich die Nachrichtenkommunikation auf ein geografisches Gebiet beschränkt und so eine gezielte forensische Untersuchung in dem jeweiligen Städtechat möglich ist. Aber da diese ihre Nutzerdaten nicht freizugänglich zur Verfügung stellen, wobei hingegen bei Twitter alle Nachrichten öffentlich und für jeden einsehbar sind, auch ohne Zugangsberechtigung zu Twitter, erübrigt sich dieser Gedanke. Über aktive Nutzerzahlen in Deutschland gibt Twitter keine Auskunft, wodurch nur Schätzungen von verschiedenen Umfragen (u.a. ARD/ZDF und BITKOM) möglich sind. Laut dem Social Media Atlas 2013 nutzen 11 Millionen Deutsche den Microbloggingdienst, was sich deutlich unter der Nutzerzahl von Facebook befindet [BUG14]. Wie schon beschrieben ist Facebook deutschlandweit das größte soziale Netzwerk, global betrachtet hat Twitter die meisten Mitglieder (1,6 Milliarden). In der Bundesrepublik sind die sozialen Netzwerke „Stayfriends“, „Wer kennt wen“ und „Xing“ hinter Facebook und Twitter ebenfalls sehr beliebt. Bundesweit haben sich 78% der Internetuser bei einem sozialen Netzwerk angemeldet (30% nutzen Twitter), was den Nutzen des TwitterSearchers sehr limitiert, da er vorerst nur auf Twitter beschränkt ist [N24] und sinnvolle forensische Untersuchungen nur mit Erweiterungen zu andern sozialen Netzwerke denkbar wären. Gründe warum Twitter in Deutschland nicht die Popularität von Facebook erreicht, lassen sich schwer definieren. Ein Problem könnte die geringere Aktivität der Nutzer sein. So wird die Plattform noch zu wenig als Kommunikationsmedium (10% der User produzieren 90% des gesamten Inhaltes)

verwendet. Auch die Beschränkung von 140 Zeichen pro Tweet bietet keinen großen Platz zur schriftlichen Entfaltung. Ein Problem mit dem Twitter früher zu kämpfen hatte war die immense Verbreitung von Spams, geschuldet der extrem frei zugänglichen API. Durch die Anpassung der API konnte aber diesem Problem entgegengewirkt werden. Auch mangelnde Feedbacks scheinen sich problematisch auf die Entwicklung von Twitter in Deutschland ausgewirkt zu haben, da viele User die Lust am twittern verloren haben, nachdem erhoffte Feedbacks ausblieben. Dieser Fehler kann aber auch bei den Usern selbst gesucht werden, da dies nicht Sinn von Twitter ist und dadurch ein mangelndes Bekanntmachen seitens der User über das Programm vorliegt. Bei Facebook werden Bilder oder Kommentare schnell mit einem Feedback versehen, wobei hingegen bei Twitter der betreffende Tweet weitergeleitet wird (Retweeted). Die größte Popularität hat Twitter in den USA. Prominente, die weltweite Bekanntheit haben, posten hier Aktivitäten für ihre Fans. Für Fernsehsendungen gibt es spezielle Hashtags, damit Zuschauer während der Sendung über diese sprechen können. Dieses Verhalten hat sich in Deutschland noch nicht etabliert. Letztendlich bietet Twitter in Deutschland bisher nur im Marketing gute Ansätze (z.B. informiert die Deutsche Bahn ihre Reisenden über Zugausfälle, -verspätungen oder andere technische Probleme mit den jeweiligen Zugverbindungen). Im privaten Sektor liegt Facebook deutlich vorn, was sich in den nächsten Jahren sicherlich auch nicht ändern wird [YEE13].

3.2.1 Anmeldung bei Twitter

Der erste Schritt, um Twitter erfolgreich zu nutzen, ist die Anmeldung, wodurch ein Account erstellt wird. Für die Anmeldung ist der vollständige Name, eine E-Mail Adresse sowie ein Passwort erforderlich, welches bei der ersten Registrierung angegeben werden muss. Nachdem man diese Parameter auf der Startseite eingetragen hat, kommt man in die nächste Anmeldemaske. In dieser werden alle angegebenen Daten überprüft, besonders die richtige Schreibweise der E-Mail-Adresse (@googlemail.com statt @gogglemail.com) und das Passwort, welches mindestens 6 Zeichen enthalten muss, wird auf seine „Stärke“ überprüft (Balken hinter dem verschlüsseltem Passwort). Danach folgt die Wahl des Benutzernamens. Twitter bietet aus dem eingegebenen vollständigen Namen einige Optionen an, welche aber nicht genommen werden müssen. Die Eingabe eines individuellen Benutzernamens ist ohne Probleme möglich. Nach der Anmeldung gibt es eine kurze Erläuterung zum Thema Tweet sowie die Aufforderung, 5 Personen beziehungsweise Organisationen zu folgen, um eine Timeline zu erstellen. Nachdem der Bestätigungslink an die angegebene E-Mail-Adresse versendet wurde, mit dem der Account bestätigt wird, steht Twitter dem Nutzer mit vollem Zugang zur Verfügung.

Das Löschen des Twitteraccounts geschieht über das Zahnrad-Symbol. In den Einstellungen findet man ganz unten die Schaltfläche „Meinen Account deaktivieren“, nach dem Betätigen kommt nochmal eine Passwortabfrage und eine nochmalige Abfrage, ob der Account wirklich gelöscht werden soll. Nach der Deaktivierung werden die Nutzerdaten noch 30 Tage gespeichert, bevor sie dauerhaft gelöscht werden. Innerhalb dieser 30 Tage ist es jederzeit möglich seinen Account zu reaktivieren. Es kann vorkommen dass auch nach der Deaktivierung Inhalte weiterhin auf Twitter sichtbar bleiben [TWI14].

Um den Twitteraccount eines Verstorbenen zu löschen müssen eine Vielzahl von Dokumenten (Kopie des Totenscheins, Anschreiben mit Bitte zu Löschung, notariell beglaubigtes Dokument mit Kontaktdaten sowie eine Kopie des Personalausweises/Führerscheins) an den Twitterhauptsitz nach San Francisco (USA) verschickt werden [WUN13].

3.2.2 Twitter-Suchmaschine

Um schnell nach Benutzern, Hashtags oder speziellen Themen in Tweets zu suchen bietet Twitter direkt eine eigene Suchfunktion an. Zum einen befindet sie sich in der oberen Menüleiste und kann, egal wo man sich zurzeit bei Twitter aufhält, jederzeit benutzt werden. Alternativ ist sie über die URL: <http://search.twitter.com> aufrufbar und zeigt zusätzlich noch sogenannte verlinkte Trending Topics an, also Begriffe (auch Hashtags) die aktuell bzw. in den letzten Stunden von vielen Nutzern verwendet wurden.



Abbildung 11 Oberfläche der Twitter-Suche

Nachteilig wirkt sich aus, dass die Twitter-Suche nicht den gesamten Datenbestand durchsucht, sondern nur Tweets, die 7 Tage zurückliegen. Abhilfe verschafft der amerikanische Social Web-Analyzer Topsy (topsy.com), der Tweets auch über den Zeitraum einer Woche suchen kann und deren Auftreten über eine gewisse Zeitspanne grafisch darstellt. Nachdem eine Suche gestartet wurde, ist es möglich über den Button „Speichern“, der sich oben in der Ergebnisliste die Suchanfrage befindet, diese für spätere Verwendungen zu sichern. Zum Löschen muss die Suchanfrage erneut aufgerufen werden und der Button „Entfernen“ (ersetzt den Speicher-Button) drücken. Über die erweiterte Suche lässt sich die eigene Suchanfrage noch konkretisieren hinsichtlich Sprache, genauen Sätzen, nach nicht zu suchenden Wörtern (Blacklist), Orte und Personen u.a. Die Suche nach Stimmungen beläuft sich ausschließlich auf in Tweets enthaltene Smileys (☺, ☹). Auch nach der Suche lässt sich eine Filterung bezüglich Personen, Fotos, Videos und Neuigkeiten differenzieren. Eine wichtige Eigenschaft ist die automatische Aktualisierung innerhalb weniger Minuten, diese erweist sich als besonders nützlich bei stark frequentierten Themen oder Sportveranstaltungen, die gerade live stattfinden. Letztendlich kann man feststellen, dass die Twitter-Suchmaschine viele Funktionen unseres TwitterSearchers erfüllt und sich ebenfalls

als wichtiges forensisches Werkzeug mit den gegebenen Funktionen im Bereich Twitter erweisen kann.

3.2.3 Zugriffsmöglichkeiten auf Twitter

3.2.3.1 Twitter-API

Die Twitter-API ist, wie schon mehrmals angesprochen, eine Programmierschnittstelle, die Drittentwicklern von den Twitter-Entwicklern zur Verfügung gestellt wurde um eine Interaktion mit dem Programm durchzuführen. Sie gestattet die Entwicklung von Zusatzprogrammen, die Twitter um Funktionen erweitert, ohne das eigentliche Programm beziehungsweise die Software zu manipulieren. Nachteilig wirkt sich aus, dass die API nur eine begrenzte Anzahl von Aufrufen zulässt. Sollte dieser Wert erreicht sein, wird man meistens ohne Vorankündigung für den Rest der angebrochenen Stunde gesperrt [SIM10].

Hierbei gibt es 3 verschiedene API-Möglichkeiten, um Zugriff auf die Twitterdaten zu erhalten:

- **Search-API** [TWI13]:

Die Search-API von Twitter ist Bestandteil der V1.1 REST-API. Sie umfasst Abfragen über Twitterdaten mittels einer Suche beziehungsweise nach einem konkreten Benutzer. Benutzer der Search-API haben die Möglichkeit, Anfragen an bereits vorhandene Tweets zu stellen, die bestimmte Suchkriterien aufweisen (z.B.: Schlüsselwörter (Keywords), Benutzernamen, Standorte u.a.). Eine einfache Art und Weise sich die Search-API vorzustellen, ist, wenn ein Benutzer eine Suche durch die Eingabe bestimmte Schlüsselwörter direkt bei Twitter (search.twitter.com) tätigen würde. Nachteilig wirkt sich aus, dass die Search-API ihren Fokus auf Relevanz richtet und nicht auf Vollständigkeit. Das bedeutet, dass einige Tweets oder User in der Suche nicht berücksichtigt werden, sondern nur die letzten 3200 weltweit verfassten Tweets, unabhängig vom Suchkriterium (mit einem speziellen Kennwort sind die letzten 5000 verfassten Tweets möglich). Einschränkung gibt es auch in der Anzahl der Anfragen. So sind zurzeit nur 180 Anfragen je 15 Minuten möglich.

Beispiele für Abfrageoperatoren (modifizierbar):

OPERATOR	Findet Tweets ...
watching now	die „watching“ und „now“ beinhalten; Standardoperator
„happy hour“	die den genauen Ausdruck „happy hour“ beinhalten
love OR hate	die entweder „love“ oder „hate“ beinhalten (oder beide)
beer –root	die „beer“ enthalten aber nicht „root“
#haiku	die den Hashtag „haiku“ beinhalten
from:helenefischer	die von der Person „helenefischer“ gesendet wurde
to:weihnachtsmann	die an die Person „weihnachtsmann“ gesendet wurde
@mashable	die die Person „mashable“ referenziert

superhero since:2010-12-27	die „superhero“ beinhalten und seit dem „2010-12-27“ gesendet wurden (Jahr-Monat-Tag)
ftw until:2010-12-27	die „ftw“ beinhalten und vor dem „2010-12-27“ gesendet wurden
movie –scary :)	die „movie“ beinhalten aber nicht „scary“ mit einer positiven Meinung
flight :(die „flight“ beinhalten mit einer negativen Meinung
traffic ?	die „traffic“ beinhalten und eine Frage stellen
hilarious filter:links	die „hilarious“ beinhalten und zu einer URL verlinken
news source:twitterfeed	die „news“ beinhalten und über TwitterFeed eingegangen sind

Darüber hinaus gibt es noch zusätzliche Parameter, die eine bessere Kontrolle der Suchergebnisse erlauben:

- *Result Type*: entscheidet ob das Ergebnis ein aktueller oder beliebter Tweet ist (oder eine Mischung aus beidem)
- *Geolocalization*: Möglichkeit eine Abfrage durch eine gegebene Position zu begrenzen (Breite, Länge, Radius) – Geocode-Parameter. Diese können sich unterscheiden zwischen statistischen Daten, also wenn Twitter auf stationären Rechnern genutzt wird oder nach mobilen Daten, wenn Twitter auf mobilen Endgeräten mit eigener Geolocation-Funktionen genutzt wird
- *Language*: der lang-Parameter beschränkt Tweets auf eine bestimmte Sprache
- **Streaming-API** [TWI12] [BRI13]:
Für das Programm TwitterSearcher wird die Streaming-API verwendet. Der Unterschied zur Search-API besteht darin, dass mit der Streaming-API Anfragen an Tweets gerichtet werden können, die sich annähernd in Echtzeit ereignet haben. Die Streaming-API bietet Entwicklern die Möglichkeit selbst vordefinierte Kriterien zu erstellen, nach denen in den Tweets gesucht wird und sobald eine Übereinstimmung gefunden wurde, wird das Ergebnis dem Entwickler ausgegeben. Aber auch die Streaming-API hat ihren Nachteil: Die gefunden Daten ähneln einer Stichprobe von Tweets, die abhängig vom Suchkriterium und dem aktuellen Datenverkehr ist. Studien haben ergeben das sich die Anzahl der gefunden Tweets zwischen 1% und 40% (niedrige Latenz) bewegen, abhängig von den oben genannten Faktoren (mit einer eingeschränkten Genehmigungsverwendung der Daten). Wenn man bedenkt, dass täglich circa 500 Millionen Tweets verfasst werden, liefert die Ergebnismenge einen verschwindend geringen Bruchteil aller Tweets weltweit, die für eine aussagekräftige Analyse jedoch nicht ausreichen. Der Grund für diese Einschränkung besteht darin, dass Twitter mit ihrer Infrastruktur es nicht unterstützen möchte, da

man den Vollzugriff auf die Tweets im Sinne des Marketings kostenpflichtig verkaufen will.

Twitter bietet verschieden Streaming Endpunkte an, benutzerspezifisch abgestimmt auf den jeweiligen Gebrauchsfall:

- *Öffentliche Streams:*
Bestehen aus öffentlichen Daten. Ideal um speziellen Usern oder Themen zu folgen beziehungsweise für Data Mining
- *Userstreams:*
Bestehen aus Streams die alle Daten eines Einzelusers beinhalten
- *Seitenstreams*
Die Multiuserversion eines Userstreams. Diese Streams sind für Server beabsichtigt, die eine große Anzahl an Usern mit Twitter verbinden.

3.2.3.2 Echtzeit-API Firehose

Da, wie schon beschrieben, der Zugriff mit der Search-API und der Streaming-API sehr eingeschränkt ist, stellen die Twitter-Entwickler eine weitere API zur Verfügung, die vollen Zugriff auf die Tweets zulässt. Die Twitter Firehose-API ist der Streaming-API sehr ähnlich, da sie ebenfalls Daten dem Endbenutzer liefert, die unmittelbar in Echtzeit erstellt wurden, mit dem Unterschied dass sie alle Tweets findet, auf die das vordefinierte Suchkriterium zutrifft. Im Gegensatz zur frei verwendbaren Streaming-API ist die Firehose-API kostenpflichtig, wodurch viele Benutzungseinschränkungen, die durch Twitter auferlegt wurden, entfallen. Diese API wird von Twitter direkt aber nur bestimmten Unternehmen angeboten wie zum Beispiel Microsoft oder Google [BRI13].

Endbenutzer, die Twitterdaten für einfache statistische Analysen verwenden wollen, können bedenkenlos die Search-API und die Streaming-API verwenden (z.B. Marketingunternehmen oder Statistiker zur Erstellung und Analyse von Trends in sozialen Netzwerken). Dies reicht aber nicht mehr aus, wenn man Abhandlungen nicht anhand von Stichproben abhängig machen kann, z.B. während eines speziellen Ereignisses (z.B. werden oft bei großen Sportveranstaltungen die aktuellen Twitterrends auf dem Fernseh Bildschirm eingeblendet, wenn sie sich auf das jeweils ausgestrahlte Sportereignis beziehen) oder eines kritischen Zustandes. Beispielsweise nutzen US-amerikanische Profisportmannschaften den Dialog in Twitter, um die Sicherheit der Zuschauer zu gewährleisten, da es mitunter schwierig ist, alles zu sehen was am Veranstaltungsort passiert. Des Weiteren ist der Echtzeitvollzugang mittlerweile in der Strafverfolgung sehr nützlich geworden, da die Polizei jederzeit weiß, was und wann in ihrem Zuständigkeitsbereich stattfindet, wobei ein paar Stichproben, wie sie die Search-API und die Streaming-API liefern nicht mehr ausreichen. Diese Beispiele finden in Deutschland aber eher wenig Verwendung und werden bevorzugt in den USA angewendet [BRI13].

3.2.3.3 Zugriffe über Reseller

Die Firehose-API wird von den Datenproviderfirmen GNIP und DataSift seit 2009 kostenpflichtig zur Verfügung gestellt, die enge Geschäftsbeziehungen zu Twitter aber auch anderen Betreibern sozialer Netzwerke führen. Ähnlich wie bei der Streaming-API wird hierbei eine Vereinbarung zwischen dem API-Endnutzer und den Anbietern der Firehose-API getroffen, auf welche Tweets der Endnutzer Zugang erhält [BRI13].

DataSift, Marktführer auf dem Gebiet (jeden Tag werden Millionen von Nachrichten verarbeitet wodurch auch der Datenspeicher täglich um mehrere Terabytes anwächst), stellt Kunden sowohl den Echtzeit- wie auch den historischen Zugang zu Twitter und anderen Daten von sozialen Netzwerken bereit, indem sie diese sammeln, nach bestimmten Kriterien filtern (z.B. Geodaten, Meinungen zu Marktforschungszwecken u.a.) und für die Bedürfnisse von Unternehmen bereitstellen. Diese Daten können zum Einen über den Erwerb der Firehose-API für die Entwicklung eigener Software eingebunden werden oder zum Anderen von DataSift selbst zusammengetragen werden, was aber sehr viel teurer ist. Für die zum Teil komplexen Filteranfragen verwendet DataSift ihre eigene Programmiersprache namens CSDL. Des Weiteren ist eine Sprachverarbeitung für Gefühlsanalysen und Entitätsgewinnung (in Kooperation mit der amerikanischen Firma Kloud) möglich, sowie die Erfassung von demographischen Daten. Um den Echtzeitzugang für einen Testbetrieb zu nutzen, ist es möglich sich, bei DataSift anzumelden und diesen für 2 Wochen kostenlos in Anspruch zu nehmen. Die Bezahlung erfolgt ähnlich einem Prepaidhandy: es erfolgt eine Überweisung eines Betrages im Voraus und sollte dieser aufgebraucht sein, so wird auch der Stream gestoppt, der die angeforderten Daten liefert. Die Gesamtkosten belaufen sich zum Einen auf den von einer DPU gemessenen Verbrauch von Rechnerkapazitäten (\$0.20 pro Stunde wobei die Minimumgebührenrate sich immer auf eine Stunde beläuft) und den Umfang der zurückgelieferten Daten (jeder gefundene Tweet kostet \$0.0001). Ebenfalls wie der Echtzeitzugang ist auch der Zugang zu historischen Daten kostenpflichtig. Dafür entwickelte DataSift den eigenen Dienst „Historics“, womit sich Tweets bis zum Januar 2010 zurückverfolgen lassen. Diese werden von Twitter in einem Archiv bereitgestellt, aufbereitet und für interessierte Unternehmen von DataSift gewinnbringend vermarktet. Ein Abonnement beginnt bei einem Preis von 3000\$ pro Monat verbunden mit einer jährlichen Verpflichtung. Inwieweit der Schutz der persönlichen Daten damit verletzt wird, sei vorerst dahingestellt. Preisnachlässe für akademische Forschungsprojekte wie den TwitterSearcher werden nicht gewährt. DataSift bietet je nach Endnutzer verschiedene API-Lösungen für private Entwickler, Organisationen und Unternehmen an [DAT14].

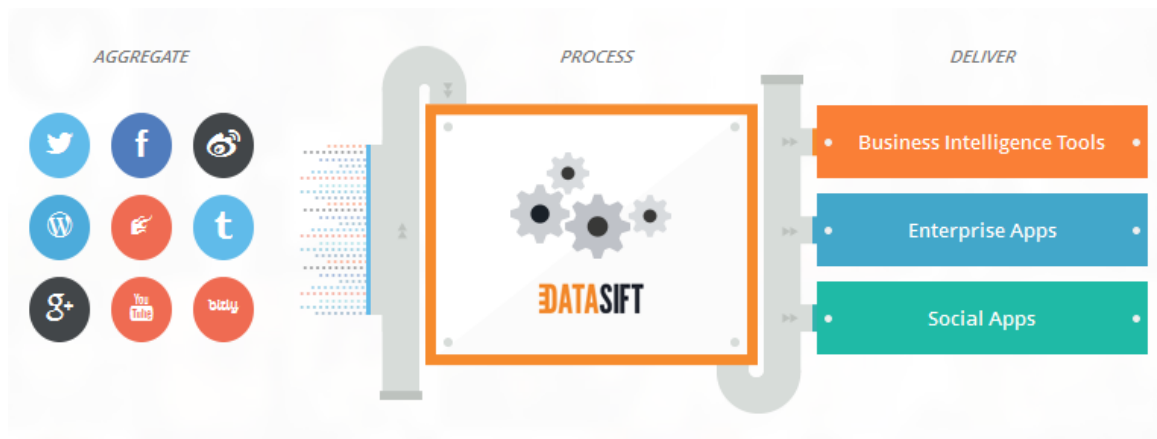


Abbildung 12 schemenhafte Funktionsweise von DataSift [DAT14]

Über GNIP hingegen bietet Twitter Datenpakete an, die zum Einen nur die Hälfte der über die Firehose-API gefunden Tweets wiedergibt (Halfhose) beziehungsweise ein Zehntel (Decahose) oder ein spezielles Datenpaket (Mentionhose), das alle Tweets enthält in denen der gesuchte Nutzer enthalten ist. Mittlerweile verfügt GNIP über ein Produkt namens Power Track, welches ebenfalls Tweets über die Firehose-API bereitstellt und somit DataSift ähnlich ist. Letztendlich bereitet GNIP ebenfalls gesammelte Daten sozialer Netzwerke (über 40) für interessierte Unternehmen auf, indem sie sie in ein gemeinsames Format (Activity Streams) umwandeln, diesen Stream mit relevanten Metadaten versehen und diesen vermarkten. GNIP soll für kleinere Firmen beziehungsweise Unternehmen bereitgestellt werden, die lediglich Studien über die Kommunikation bei Twitter erstellen wollen. Ebenso wie DataSift bietet GNIP Zugang zu Echtzeitdaten und historischen Daten [GNI14].

3.2.4 Charakteristika von Tweets

Wie schon des Öfteren erwähnt, läuft die Kommunikation bei Twitter über sogenannte Tweets ab. Diese Kurznachrichten haben eine maximale Zeichenlänge von 140. Jene Zahl steht beim Verfassen eines Tweets unter dem Eingabefenster und wird heruntergezählt sobald mit dem Schreiben des Tweets begonnen wird. Wird die Zeichenanzahl überschritten, zeigt Twitter dies mit roten negativen Zahlen an und verhindert somit, dass der Tweet gesendet werden kann. Sobald die Nachricht fertig geschrieben ist, kann über den Button „Twittern“ die Nachricht versendet werden [SIM10].

Ein Tweet besteht aus einem Profilbild (Avatar) und dient in der Regel zur Erleichterung der optischen Unterscheidung der jeweiligen Twitternutzer. Jedes Profilbild ist automatisch auf das dazugehörige öffentliche Profil des jeweiligen Nutzers verlinkt. Des Weiteren steht zu Beginn eines Tweets in der Regel der Name des Twitternutzers (insofern er/sie seinen richtigen Namen angegeben hat), der ebenfalls mit dem jeweiligen Profil verlinkt ist und der Twittername. Dieser Name beginnt mit einem @-Zeichen und bedeutet hinter dem Benutzername, dass der Tweet eine Antwort an einen anderen Twitternutzer enthält [SIM10]. Die Tradition dieses Zeichens, welches wie das englische at (deutsch: an bzw. bei) gesprochen wird, im Internet besonders in Chats zu verwenden, ist nicht die Erfindung von Twitter, sondern diente in Chaträumen zur Erhaltung der Übersicht, denn mit dem @-Zeichen vorangestellt konnte gezielt der jeweilige Kommunikationspartner angeschrieben

werden. Während ein normaler Tweet eine ungerichtete Nachricht an die ganze Onlinewelt ist, kann ein Tweet mit einem „@Username“ an einen bestimmten User direkt adressiert werden, welcher jedoch immer noch für das ganze Twitteruniversum sichtbar ist [SIM10]. Neben dem Benutzernamen steht noch die Zeitangabe mit dem Alter des Tweets, diese Zeit kann in Sekunden, Minuten oder Stunden angegeben werden. Tweets die älter als 24 Stunden sind, werden mit dem Erstellungsdatum versehen. Das genaue Erstellungsdatum kann man aber jederzeit abrufen, indem man mit dem Cursor über das Datum fährt und dieses nach kurzer Zeit mit Uhrzeit und Datum angezeigt wird. Twitter verwendet dabei das amerikanische System AM und PM und lässt (noch) keine 24 Stunden Anzeige zu (Beispiel: Tweet wird um 18:30Uhr geschrieben – Twitter zeigt 6:30 PM an). Damit auch die Erstellungszeit fremder Tweets richtig angezeigt wird, muss der Twitter-Account auf die lokale Zeitzone des Nutzers eingestellt sein [SIM10]. Nach diesen Angaben folgt die eigentliche Nachricht. Tweets können reine Textnachrichten sein, können Bilder enthalten, die man über das Kamerasymbol hinzufügen kann, Video- und Audiodateien (über spezielle Programme: „Twitter Player Cards“) beinhalten, URLs oder sogenannte Hashtags. Diese Textanhängsel mit zusätzlichen Metainformationen werden durch ein vorangestelltes #-Zeichen (englisch: hash) gekennzeichnet, gefolgt von einem Schlagwort oder einer alphanumerischen Abkürzung. Über gemeinsam gebloggte Hashtags lässt sich eine gewisse Gemeinschaft erkennen, zum Beispiel Nutzer die das gleiche Thema berührt, Teilnehmer an bestimmten Veranstaltungen sind oder gemeinsame Interessen teilen (Sportvereine, TV-Serien u.a.). Hashtags sind in Tweets direkt anklickbar und nutzen die Twitter-Suche um Tweets mit dem benutzten Hashtag zu finden. Organisatoren von Veranstaltungen geben offizielle Hashtags heraus, um sicherzustellen dass Tweets zuverlässig gefunden werden [SIM10]. Nachteilig wirkt sich aus, dass viele Hashtags innerhalb eines Tweets diesen sehr unübersichtlich machen. Des Weiteren gibt es unter der Textnachricht noch verschiedene Symbole. Dazu zählen „Antworten“ für Kommentare oder Meinungen Dritter, die wiederum vom Verfasser auf dieselbe Weise kommentiert werden können. Ferner zählt das Mülltonnen-Symbol („Löschen“) dazu. Sollten Tweets Rechtschreibfehler enthalten, so lassen sich diese nicht mehr ausbessern. Die einzige Möglichkeit besteht darin den kompletten Tweet über das Mülltonnen-Symbol zu löschen. Damit verschwindet der Tweet zwar aus der Zeitleiste des Nutzers, ist aber über die Twitter-Suche weiterhin auffindbar. Ob er physisch bei Twitter gelöscht wurde, kann nie mit genauer Wahrscheinlichkeit gesagt werden, weiterhin kann er bereits retweeted worden sein und so weiterhin bei Twitter vorhanden sein [SIM10]. Das Mülltonnen-Symbol erscheint jedoch nur bei selbst verfassten Tweets, da man ausschließlich diese löschen kann. Bei fremden Tweets findet man alternativ an dieser Stelle die Retweet-Funktion, welche eine komplette Kopie des Tweets für die eigene Zeitleiste erstellt. Das nächste Symbol in der Leiste ist ein Stern („Favoriten“), der eine ähnliche Funktion wie die Lesezeichen bei Browsern hat. Damit lässt sich ein Tweet markieren, den man sich noch einmal ansehen oder merken möchte. Eine Auflistung aller Favoriten (Favs) kann man sich auf der eigenen Startseite unter dem Reiter „Favoriten“ ansehen [SIM10]. Neben diesen Symbolen ist es noch möglich über den Link „Mehr“ → „Tweet einbetten“ den verfassten Tweet zum eigenen Artikel oder zur eigenen Website hinzuzufügen. Dafür stellt Twitter den Quellcode des Tweets dem Nutzer zur Verfügung. Der Link „Öffnen“ ermöglicht eine weitere Detailansicht, in welcher man ebenfalls das

Erstellungsdatum einsehen kann sowie wer meinen Beitrag retweeted hat und zu seinen Favoriten hinzugefügt hat mit den dazugehörigen Profilbildern.



Abbildung 13 Beispieltweet mit den wichtigsten Bestandteilen

Einer Studie von 2012 zufolge, finden Tweets und Links in Tweets die bestimmte Eigenschaften aufweisen, mehr Beachtung als andere. Dazu zählen [ZAR12]:

- Tweets mit einer Zeichenlänge von 120-130
- Links die im ersten Viertel des Textes platziert wurden
- je häufiger in einer Stunde getwittert wird, umso weniger werden die Links angeklickt
- Links/Tweets die die Begriffe „Via“, „@“, „RT“, „Please“ oder „Check“ enthalten
- Links/Tweets die aus Verben und Adverbien bestehen werden häufiger angeklickt als welche die fast nur aus Nomen und Adjektiven bestehen
- Links/Tweets die am Wochenende verfasst wurde
- Links/Tweets die am späten Nachmittag geschrieben wurden

4 Forensische Ansätze

4.1 Forensische Informatik

Das Wort „forensisch“ entstammt dem lateinischen Wort *forum* was übersetzt Marktplatz bedeutet, dem Ort an dem früher die Gerichtsverhandlungen abgehalten wurden. Somit bedeutet Forensik „die Anwendung wissenschaftlicher Methoden auf Fragen des Rechtssystems z.B. zur Untersuchung und Verfolgung von Straftaten“ [DEW11 Seite 5]. Forensische Wissenschaften beschränken sich aber nicht nur auf den digitalen Bereich, sondern finden sich auch in der Psychiatrie, Medizin oder Linguistik.

Seitdem die Computertechnik in immer mehr Bereiche des öffentlichen und geschäftlichen Lebens Einzug gehalten hat und sich soziale Netzwerke immer größerer Beliebtheit erfreuen, stieg auch die Anzahl von Straftaten über digitale Geräte und soziale Plattformen. In der Regel befasst sich die Forensische Informatik mit der Entstehung, Erkennung, Bewertung und Sicherung digitaler Spuren. Besonderes Augenmerk gilt dabei der Sicherung und Verwertung der digitalen Spuren, da diese gerichtsfest sein müssen, so auch die Daten, die der TwitterSearcher letztendlich auf die gestellte Suchanfrage liefert [DEW11].

4.2 Kriminalität in sozialen Netzwerken

Der Gedanke, dass Menschen Straftaten in sozialen Netzwerken zugeben oder ankündigen, klingt auf den ersten Blick sehr unwahrscheinlich, dennoch gibt es dokumentierte Beispiele aus den USA und Kolumbien bei denen dies dennoch der Fall war. In einem der beiden Beispiele, welches sich im Jahr 2012 ereignete, kündigte ein Twitter-Nutzer in mehrere Kurznachrichten einen Anschlag auf ein Theater in New York City an. Twitter weigerte sich jedoch mit der Herausgabe des Benutzernamens, der E-Mail-Adresse sowie den Log-Daten, die unter anderem auch den Standort des Nutzers beim Versenden des Tweets enthalten, an die Behörden. Twitter ist generell nicht dazu verpflichtet Nutzerdaten öffentlich darzulegen, obwohl es im geschilderten Fall wichtig gewesen wäre. Warum die Polizeibehörden nicht auf einen der zuvor erläuterten Wege die Informationen zusammengetragen haben ist nicht erwähnt. In den USA können Gerichte Unternehmen zur Herausgabe von Nutzerdaten zwingen „wenn sich Gründe für eine strafrechtliche Ermittlung ergeben“. Ebenso ist die Übermittlung von Daten in andere Länder möglich. Dies erweist sich als vorteilhaft, da sich die Betreiber sozialer Netzwerke meistens im Ausland befinden und so eine Freigabe der Benutzerdaten erheblich erleichtert wird [FAZ12]. Das zweite Beispiel ereignete sich konkret bei Facebook, wäre aber auch ein Szenario, welches sich bei Twitter zugetragen haben könnte. Dabei fand die kolumbianische Polizei 2010 zwei Leichen von Jugendlichen ohne einen Hinweis auf den Mörder beziehungsweise das Motiv. Wenige Tage vor der Tat wurde eine Liste von 69 jungen Männern bei Facebook gepostet auf der sich auch die beiden Opfer befanden. Fünf Tage später stirbt ein weiterer Jugendlicher der ebenfalls auf der Liste stand. Nach den drei Morden befinden sich plötzlich 31 Namen mehr auf der besagten Liste und

mehrere Drohungen. Der Fall ist bis heute nicht aufgeklärt worden, auch weil sich die Spuren, die zu den Tätern führen könnten, im Facebookuniversum verlieren [WDW12]. Dieser Fall macht deutlich, dass trotz moderner forensischer Verfahren sich nicht jede Straftat aufklären lässt, weil sich die Täter die Anonymität innerhalb des Netzwerkes zu nutze machen.

Die genannten Beispiele sind sicherlich Extremfälle und werden in dieser Art und Weise eher selten vorkommen. Wahrscheinlicher ist es da, dass Nutzer sozialer Netzwerke als Zeugen fungieren und ihre Beobachtungen eventuell sogar mit Bildern oder Videos online stellen und so das Interesse der Ermittler wecken. Alternativ kann man aber auch soziale Netzwerke dazu verwenden, gestohlene oder vermisste Gegenstände bzw. Personen wiederzufinden (Beispiel: Foto von einem gestohlenen Fahrrad mit der Überschrift „Wer hat es zufällig gesehen“). Solche Nachrichten können entscheidende Hinweise in den Antworten enthalten, die für den Verfasser mitunter nicht weiterverfolgt werden können, für die Kriminalisten aber eine verfolgbare Spur liefern können. Es zeigt sich auch, dass immer mehr Ermittler soziale Netzwerke benutzen um z.B. Fahndungsbilder zu veröffentlichen, die sich aufgrund der großen Nutzergemeinde schnell verbreiten. Besonders bei Twitter können Hastags wichtige Informationen auf das Umfeld und die Interessen von Tätern oder Zeugen geben.

Eine große Gefahr, die ebenfalls mit sozialen Netzwerken einhergeht, besteht darin, dass Täter gezielt Nachrichten nach Lebensgewohnheiten durchsuchen. Reiseunternehmungen sollten daher nicht gepostet werden oder auf ein Minimum beschränkt werden, da dies Einbrechern Anlass für einen Diebstahl geben könnte. Niederländische Programmierer wollen mit der Webseite *pleaserobme.com* auf das Risiko solcher Nachrichten aufmerksam machen. Eine von ihnen entwickelte Software filtert Nachrichten nach unbeaufsichtigten Wohnungen und verknüpft diese mit Bildern von Gebäuden oder Fahrzeugen. Nachteilig kann sich auch die Geotagging-Funktion dabei auswirken, die unter anderem bei Twitter oder Facebook vorhanden ist. Diese gibt Auskunft wo sich der Nutzer gegenwärtig, während des Verfassens seiner Nachricht, befindet und kann so dem Täter Auskunft geben, wieviel Zeit er noch zum Vollenden seines Deliktes hat [WDW10].

Soziale Netzwerke bieten Tätern leider die Möglichkeit, eine Vielzahl von Straftaten zu begehen, seien es Cybermobbing, Spionage oder Vergehen (Mord, Erpressung, Betrug u.a.), indem man sich ein Fakeprofil erstellt und damit ahnungslose Menschen in eine Falle lockt. Auch das Phishing, also gezielt falsche Profile oder E-Mails zu erstellen, um an Nutzerdaten heranzukommen, tritt immer häufiger auf. Um die Straftaten, die in sozialen Netzwerken begangen werden, aufzuklären brauchen Kriminalisten immer leistungsstärkere forensische Software, da die Menge an Daten unüberschaubare Ausmaße angenommen hat.

4.3 Computerlinguistische Analyse von Twitterdaten

Die Computerlinguistik beschäftigt sich im Allgemeinen mit der maschinellen Verarbeitung natürlicher Sprache [UVT14].

Wie schon mehrfach erwähnt wurde, liefern sowohl die integrierte Twittersuche als auch Suchsysteme, die auf der Streaming-API basieren, sehr unvollständige Ergebnisse. Eine Lösung bietet dabei die angesprochene Firehose-API, jedoch sollte zu Bedenken sein, ob der enorme finanzielle Aufwand dem Zweck der Informationsbeschaffung gerechtfertigt ist. Auf Grund der Unvollständigkeit der gefundenen Twitterdaten gestaltet sich eine computerlinguistische Analyse als sehr oberflächlich. Dennoch können einige interessante linguistische Fakten bezüglich Twitter vermerkt werden. Die meistgeschriebene Sprache auf Twitter ist natürlich Englisch, gefolgt von Japanisch, Portugiesisch, Indonesisch und Spanisch. Der Grund, warum diese Sprachen bevorzugt bei Twitter verwendet werden, liegt in ihrer globalen Verbreitung. Da in der Regel die meisten gefundenen Tweets durch den TwitterSearcher englischsprachig sind, würde dieser Aspekt die Analyse durch die Kriminalisten sehr erschweren. Auf Grund dessen verfügt die Anwendung über eine Sprachauswahl. Auffällig in dieser Auflistung ist zum einen der hohe Anteil an Indonesisch, welcher durch die große Beliebtheit des Microbloggingdienstes auf dem Inselstaat begründet ist. Zum anderen fällt auf, dass die meistgesprochene Sprache der Welt, das Hochchinesisch, nicht im Ranking enthalten ist. Das hat den Hintergrund, dass der Microbloggingdienst in der Volksrepublik China seit 2009 blockiert wurde.

Die Analyse des Textkorpus von Twitter beinhaltet viele Elemente die im Kapitel Charakteristika von Tweets angesprochen wurden: #Hashtags, URLs aber auch die Verwendung von Emoticons (einzige Möglichkeit der Twittersuche, auf Stimmungen des Tweetverfassers zu schließen). Des Weiteren gilt es Ausdrücke der Umgangssprache, von Slangs oder Dialekten richtig einzustufen und zu bewerten. Besonderes Augenmerk gilt dabei auch der Prüfung von Tippfehlern und fremdsprachigen Homogrammen (Wort mit der gleichen Schreibweise aber unterschiedliche Bedeutung z.B. das deutsche Wort „war“ kann mit dem englischen Wort „war“ (Krieg) verwechselt werden) [SCH13]. Die Ausnutzung der geschilderten Eigenschaften kann manuell oder softwarebasiert erfolgen, um Muster in der Sprache oder eine Autorenbestimmung vorzunehmen.

4.4 Emotionserkennung in sozialen Netzwerken

Wie beschrieben, beläuft sich die Emotionserkennung in ihrer herkömmlichen Form nur auf das Finden von Emoticons. Natürlich kann ein Twitternutzer seine Gefühle auch in schriftlicher Form ausdrücken, er wird dabei aber von der Zeichenbegrenzung eingeschränkt, was sich darin äußert, dass die Emotionen weniger förmlich ausgedrückt werden. Das Interesse, seine Gefühle in sozialen Netzwerken zum Ausdruck zu bringen, hat immense Ausmaße angenommen, die eine Erforschung auf diesem Gebiet sinnvoll machen.

Amerikanische Forscher der Universität Texas haben einen Textkorpus von Tweets gesammelt und erläutern daran die Gefühle Wut, Abneigung, Angst, Freude, Traurigkeit, Überraschung (6 Basis-Gefühle nach Paul Ekman) und Liebe (da die Annahme besteht, dass diese Emotion oft bei Twitter vorkommt). Dazu benutzen sie den Textkorpus von einer Vielzahl von Twitterthemen, die die beschriebenen Emotionen hervorgerufen haben, um einen Klassifizierer zu entwickeln der über Lernmethoden automatisch Gefühle in Tweets

entdeckt. Dieses System verfolgt ebenfalls die Emotionen eines Einzelnen über eine gewisse Zeitspanne. Solche Erkenntnisse sind nicht nur für Forensiker von Nutzen, sondern auch für Betriebe, die Waren verkaufen oder Dienstleistungen anbieten, um zu beurteilen wie ihr Produkt beziehungsweise ihre Dienstleistung in der Öffentlichkeit ankommt. Diese Recherchen sind abhängig von einer großen Anzahl an Tweets und lexikalischen Entscheidungsregeln (Heuristiks), die das Ergebnis zuverlässiger machen.

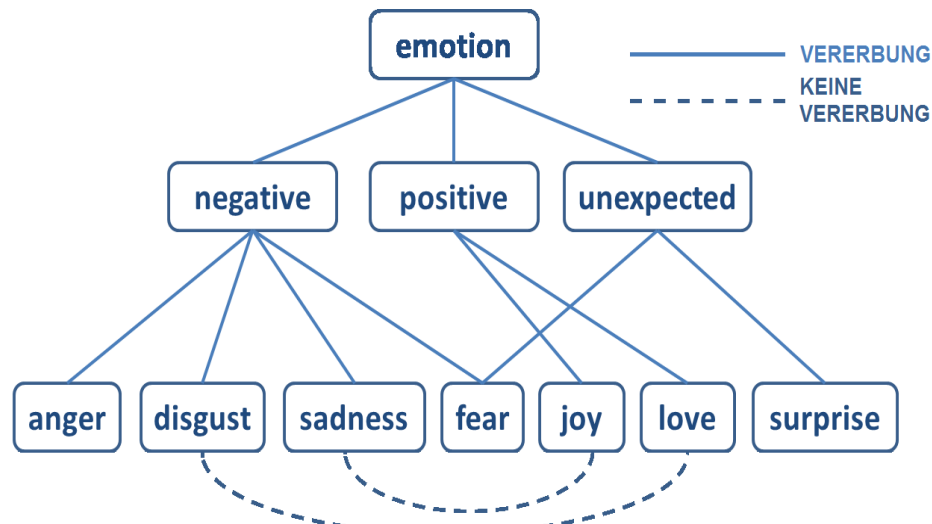


Abbildung 14 Emotionsontologie für die 6 Ekman's Basisgefühle + Liebe [ROB12]

Für jedes Thema (u.a. Super Bowl, Valentinstag, 11.September, US Wahl 2012, Weihnachten) wurde eine Liste von Hashtags kompiliert, die in den betreffenden Tweets enthalten sind, um die Themen und Trends zu kennzeichnen. Die Tweets wurden mithilfe der Twitter-API gedownloadet und durch Kommentatoren hinsichtlich der enthaltenen Gefühle (oder keine enthalten) bewertet (z.B.: Liebesnachrichten enthielten sowohl Freude als auch Traurigkeit). Die meisten analysierten Tweets dieser Studie enthielten keine Hinweise auf Gefühle. Am häufigsten anzutreffende Gefühle waren Abneigung und Freude. Die Aussagekraft ist jedoch sehr oberflächlich, da nur ein paar wenige Themen ausgewählt wurden und nur wenige Gefühle ausgewählt werden konnten. Um diese Forschungen weiter voranzutreiben, wurde auf den ersten computerbasierten Mehrzwecktextanalysierer (General Inquirer) zurückgegriffen. Basierend auf psychologischen Wörterbüchern bezogen auf Status, Motive und soziale und kulturelle Rollen. Er enthält viele unterschiedliche semantische Kategorien, die bisher unberücksichtigt blieben und ein besseres Ergebnis liefern können.

Darüber hinaus gibt es noch weitere Forschungen auf dem Gebiet der Emotionserkennungen. Nicht nur in sozialen Netzwerken sondern auch in E-Mails, Schlagzeilen von Nachrichten und Selbstmordbriefen. Andere beziehen sich ausschließlich auf das Vorkommen von Emoticons. Hinsichtlich der Verfasser von Tweets mit gleichen Gefühlen wurde eine Methode entwickelt die sich auf linguistische Analysen beruft (Linguistic Inquiry Word Count (LIWC)-Methode). Diese misst den Wortgebrauch anhand psychologisch aussagekräftiger Kategorien (Pronomen- und Verbensätze, Wörter, die positive oder

negative Gefühle ausdrücken, Wörter, die soziale Beziehungen oder Hierarchien kennzeichnen u.a.) und legt damit einen zusätzlichen Fokus auf die untersuchten Tweets [ROB12].

4.5 Softwarebasierte und analytische forensische Ansätze zur Personensuche in sozialen Netzwerken

Nachfolgend werden einige Ansätze erläutert wie man mit den vorhandenen oder gefundenen Daten beziehungsweise Informationen auf Täter oder Zeugen schlussfolgern kann. Darüber hinaus gibt es sicherlich noch mehr Ansätze, die den Rahmen dieser wissenschaftlichen Arbeit jedoch sprengen würden.

4.5.1 First Story Detection (FSD)

Dieses System wurde an der University of Edinburgh (Schottland) entwickelt mit der anfänglichen Aufgabe themengleiche Nachrichtenstreams zu ermitteln und den Ursprung, also den ersten Beitrag, der zu dem betreffenden Thema veröffentlicht wurde zu finden. Zunächst bezog sich diese Aufgabe noch nicht auf soziale Netzwerke, sondern auf Streams von Nachrichtenagenturen. Der Gedanke, das FSD auf Tweets anzuwenden, ergab sich aus der gewachsenen Beliebtheit des Microbloggingdienstes sowie den zum Teil schnelleren Ereignismeldungen, bezogen auf Nachrichtenagenturen. Die hohe Datenmenge (500 Millionen Tweets pro Tag) wirkt sich dabei negativ auf Suchzeit und -ergebnis aus, verglichen mit der Suchzeit bei Streams von Nachrichtenagenturen. Demgegenüber steht der Vorteil der sozialen Komponente. Es lässt Einblicke zu, welche Auswirkungen ein Ereignis auf die Menschen hat und wie sie damit umgehen. Der Algorithmus vergleicht alle neuen eintreffenden Nachrichten mit den bereits vorhandenen, kontrolliert das Tf-idf-Maß¹ des Dokumentes (z.B. Ähnlichkeit von Nomen) und versucht daraus verschiedene Mengen zu erstellen oder fügt neue Dokumente chronologisch einer bestehenden Menge hinzu. Spammitteilungen und andere unwichtige Nachrichten (z.B. über aktuelle Gemütslagen eines Nutzers) werden mit einer vertretbaren Genauigkeit herausgefiltert. Erste Versuche ergaben, dass sich Nachrichten von verstorbenen Prominenten am schnellsten bei Twitter verbreiten. Ein möglicher Nutzen, der sich für die Ermittler daraus ergeben könnte, wäre den Urverfasser einer Nachricht zu ermitteln und so denjenigen Nutzer zu extrahieren, der wahrscheinlich wirklich Zeuge oder Täters eines Verbrechens war und nicht diejenigen Personen die den Beitrag nur retweeted haben und zur Tat nicht beitragen können. Weiterhin könnte man über seine Bekannten bei sozialen Netzwerken auf den Urverfasser schlussfolgern oder über deren Metadaten, Hashtags oder Links, die wie bereits erwähnt fast jeder Tweet enthält [FSD10].

¹ „Maß zur Beurteilung der Relevanz von Termen in Dokumenten“ [WIK14]

4.5.2 soziale netzwerkbasierte Entitätsgewinnung für Personenontologien

Einen weiteren interessanten Ansatz liefern zwei Forscher New Yorker Universitäten, um eindeutige Personensuchergebnisse in Suchmaschinen sozialer Netzwerke zu erzielen. Bedenkt man, dass eine Suchanfrage eines einzelnen Vornamens beziehungsweise Nachnamens eine Unmenge an Treffern in sozialen Netzwerken liefert, so ist es ziemlich mühselig den eigentlich Gesuchten zu finden. Und selbst wenn man den Vor- und Nachnamen kennt, kommt es mit großer Wahrscheinlichkeit vor, dass es viele Menschen (in diesem Modell die Entitäten) mit dem selben Namen gibt. In der Informatik beschreibt eine Entität ein „bestimmtes Objekt in der Datenmodellierung über das Informationen gespeichert oder verarbeitet werden soll“ [WIK141]. Die New Yorker Forscher haben ein Modell entwickelt, um dieses Homonymproblem einzuschränken. Dafür wird eine Ontologie von Personen mit demselben Namen und deren Eigenschaft erstellt, die anhand dieser Eigenschaften in verschiedene Klassen eingeteilt werden und daran eindeutiger zu finden sind. Das System, welches auf diesen Grundlagen entwickelt wurde, trägt den Namen Ontology Supported Web Search. „Ontologien sind geordnete Darstellungen von Begriffen und den zwischen ihnen bestehenden Beziehungen. Zur Darstellung kann eine Ontologie aus verschiedenen Bestandteilen bestehen: Begriffen, Typen, Instanzen, Relationen“ u.a. [WIK142]. Anfänglich haben sie sich dabei auf Namen von berühmten Personen bezogen, die nach der Länge des Eingabestrings ihres Namens in die Suchmaschine in drei Klassen (A, B, C) eingeteilt wurden. Je weniger Buchstaben eingegeben werden mussten und je schneller die Autovervollständigung den Namen parat hatte, umso höher war die Klassifizierung. Jedoch wäre es auch denkbar, dies auf normale Personen anzuwenden, auch wenn dies aufwendiger ist als bei berühmten Menschen, da es mehr normale Menschen mit dem selben Namen gibt als Prominente. Die benötigten Eigenschaften werden aus sozialen Netzwerken und DBpedia (extrahiert Informationen aus Wikipedia für Webanwendungen) gewonnen und bereinigt. Die sozialen Netzwerke wurden dabei als Informationsquelle aufgegriffen, da sich diese schnell aktualisieren, was den großen Vorteil gegenüber DBpedia darstellt, welches zum Teil in dieser Hinsicht sehr statisch ist. Hieraus entsteht auch das Problem zur Klassifizierung normaler Menschen. Diese besitzen bei DBpedia keinen Eintrag und können somit nur über die sozialen Netzwerke klassifiziert werden. Twitter speichert dabei eine Vielzahl von Attributen des jeweiligen Nutzers (z.B.: User-ID, Name, Profilbild-URL, Standort, Follow Count, Friend Count, Zeitzone u.a.), über die ebenfalls eine Ontologie erstellt werden könnte und der Nutzer eindeutig zugeordnet werden kann [TIA13]. Der TwitterSearcher extrahiert einige diese Eigenschaften (Name, User-ID), die dann zur gezielten forensischen Personensuche verwendet werden können.

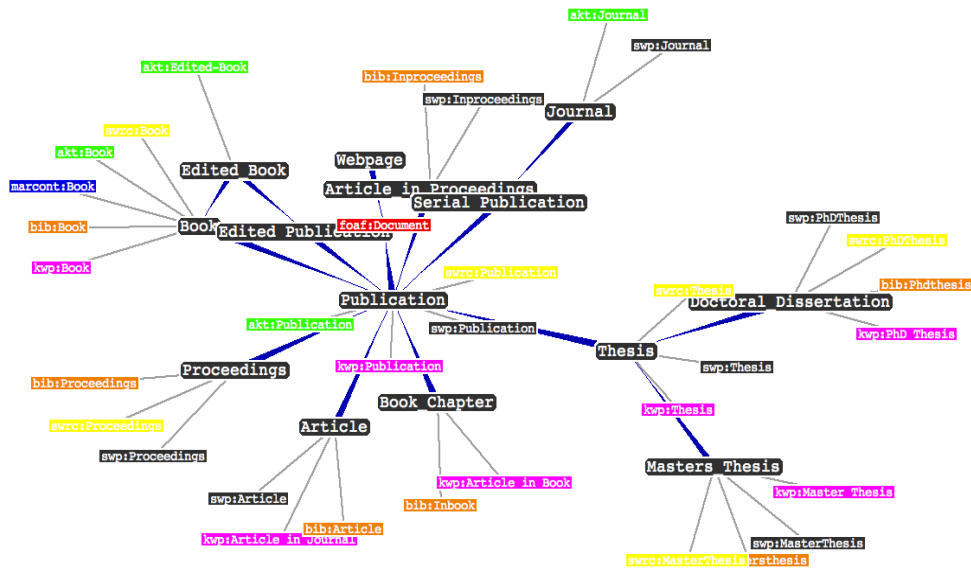


Abbildung 15 Beispiel einer einfachen Ontologie [HOL13]

4.5.3 Erzeugung von Bekanntenempfehlungen in sozialen Netzwerken durch Mining des Benutzercontent

Der Zweck eines sozialen Netzwerkes ist neben den geführten Kommunikationen auch die Vergrößerung des persönlichen Bekanntenkreises. Facebook liefert unter der Rubrik „Personen, die du vielleicht kennst“ Bekanntenvorschläge, welche durch verschiedene Attribute generiert werden. So erstellt Facebook diese Liste zum Beispiel anhand des Geburtstages, indem es Gleichaltrige darstellt oder aus angegebenen Fakten des Lebenslaufes (z.B. Schule, Wohnort) die mit den aufgelisteten Nutzern übereinstimmen oder ähnlich sind. Ebenso führen die selben Interessen (Sport, Musik, Fernsehen, Bücher u.a.) oder besuchte Nutzerseiten zu solchen Empfehlungen. Facebook gibt über diesen Algorithmus aber keine detaillierten Auskünfte, jedoch erfolgt dabei die Erforschung von neuen Bekannten anhand des sozialen Graphen, in dem alle Beziehungen zu Bekannten und deren Eigenschaften enthalten sind. Wissenschaftler der Universität Cagliari (Italien) haben ihren eigenen Algorithmus entwickelt, der jedoch nur die verschiedenen Interaktionen eines Nutzers als Grundlage nimmt und nicht den sozialen Graphen und daraus Empfehlungen von Bekannten in einem sozialen Bookmarksystem erstellt. Über diese Systeme ist es möglich, Lesezeichen von beliebten Seiten zu hinterlegen auf die mehrere Nutzer Zugriff haben. Die Lesezeichen haben dann analog die Funktion wie die Interessen bei Facebook. Diese Interaktionen werden durch Datamining gewonnen, was so viel bedeutet wie aus einer großen Datenmenge das Nützlichste zu entnehmen. Er berechnet zunächst, wie ähnlich sich die gegebenen Tags (Kategorie bzw. Schlagwort eines Lesezeichens) sind und danach die Ähnlichkeit der Bookmarks. Aus den errechneten Erkenntnissen generiert er dann die Empfehlungen für den Zielnutzer. Bei Twitter gestalten sich diese Bekanntenempfehlungen jedoch etwa anders. Da es beim Mikrobloggingdienst keine „Freunde“ in eigentliche Sinne gibt, sondern Personen, denen man folgt (Follower), lässt sich der beschriebene Algorithmus nur bedingt einsetzen und liefert ungenaue Ergebnisse. Jedoch gibt es andere wissenschaftliche Ansätze die als Grundlage den

sozialen Graphen von Twitter verwenden, in dem retweetete Tweets, gepostete Hashtags und URLs sowie Follower enthalten sind, über die ebenfalls eine Empfehlung, in Form von Followern erstellt werden kann. Für Kriminalisten ist diese Tatsache in soweit interessant, da sie über gemeinsam geteilte Interessen auf das Umfeld eines Täters oder Zeugen schließen können und so zusätzliche Informationen erhalten, die die weiteren Ermittlungen voranbringen können [MAN13].

5 TwitterSearcher

5.1 Theoretische Grundlagen

5.1.1 Eclipse

Die javabasierte Entwicklungsumgebung (IDE), in der der TwitterSearcher entwickelt und programmiert wird, trägt den Namen Eclipse. Dafür wird die objektorientierte Programmiersprache Java verwendet.

Projektname	Version	Veröffentlichung
	3.0	28. Juni 2004
	3.1	28. Juni 2005
Callisto	3.2	30. Juni 2006 ^[4]
Europa	3.3	30. Juni 2007 ^[5]
Ganymede	3.4	25. Juni 2008 ^[6]
Galileo	3.5	24. Juni 2009 ^[7]
Helios	3.6	23. Juni 2010 ^[8]
Indigo	3.7	22. Juni 2011 ^[9]
Juno	4.2	27. Juni 2012 ^[10]
Kepler	4.3	26. Juni 2013 ^[11]
Luna	4.4	angekündigt für 25. Juni 2014 ^{[12][13]} (bereits als Entwicklungsversion verfügbar)
Mars	x.x	noch kein Termin bekannt ^[14]

■ Ältere Version; nicht mehr unterstützt
 ■ Ältere Version; noch unterstützt
 ■ Aktuelle Version
 ■ Zukünftige Version

Abbildung 16 Übersicht der Eclipse-Versionen [WIK144]

Die Anfänge gehen auf das Jahr 2001 zurück, als IBM, OTI und acht weitere Firmen Eclipse der Öffentlichkeit präsentierten. Es arbeitet nach dem Open-Source Prinzip, was bedeutet, dass es nahtlos mit anderen offenen oder kommerziellen Tools der Anwendungsentwicklung zusammenarbeiten kann – eine Tool-Integrationsplattform. Sie lässt sich je nach Aufgabenstellung mit verschiedenen Plug-Ins erweitern. Des Weiteren ist Eclipse plattformunabhängig und ermöglicht, neben der bereits angesprochenen Programmiersprache Java, die Entwicklung von Software in C, C++, HTML und PHP u.a. [SHA04].

Der Java-Editor verfügt über eine Vielzahl von nützlichen Funktionen, wie der Syntaxhervorhebung, Codevervollständigung, einem Debugger, Suchwerkzeugen, die Codes über Projekte und Bibliotheken hinweg finden können, sowie eine Codeassistentz, in der Vorlagen für Codemuster enthalten sind [SHA04].

Wie bereits angesprochen basiert die Plattformarchitektur auf dem Konzept von Plug-Ins. Plug-Ins sind dabei strukturierte Code- oder Datenbündel um die Funktionalität des Programms zu erweitern. Beispiele hierfür sind Bibliotheken (Javaklassen mit einer öffentlichen API), Plattformerweiterungen oder Dokumentationen. Einige Plug-Ins erweitern die Plattform hinsichtlich optischer Eigenschaften unter Benutzung des Extension Modells. Das Eclipse-SDK beinhaltet die Basisplattform sowie zwei Werkzeuge die nützlich für das Plug-In Development sind [ECL14]:

- *Java Development Toolkit (JDT)* → implementiert die Javaentwicklungsumgebung
- *Plug-In Developer Environment (PDE)* → fügt spezielle Werkzeuge hinzu, die die Entwicklung von Plug-Ins und anderen Erweiterungen optimieren [ECL14]

Das Java Development Toolkit und das Plug-In Developer Environment sind letztendlich auch nichts anderes als Plug-Ins.

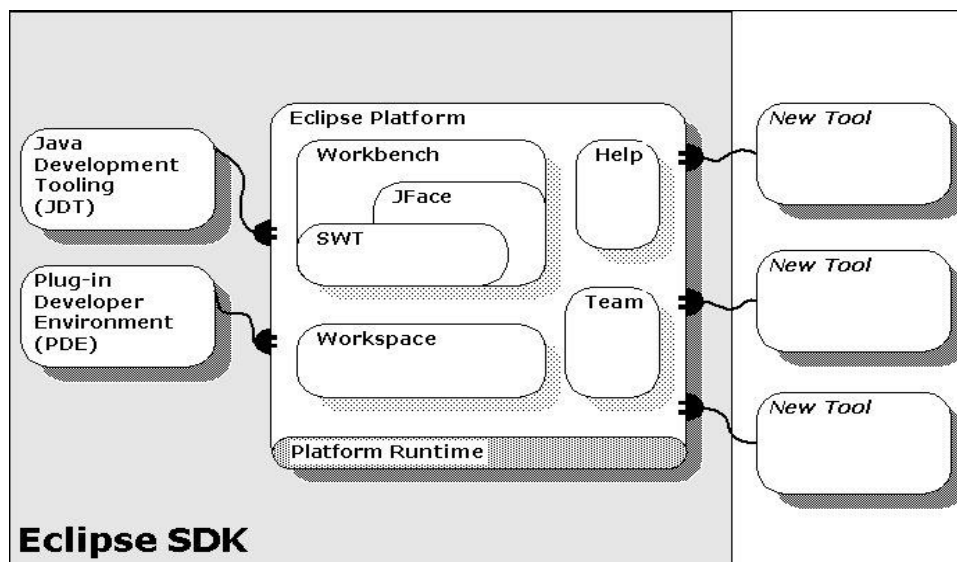


Abbildung 17 Plattformarchitektur von Eclipse [ECL14]

Die „New Tools“ in Abbildung 5.2. stellen neue Plug-Ins dar, von denen sich beliebig viele in Bezug auf Anzahl und Hersteller integrieren lassen. Der Workspace ist die Arbeitsumgebung, in der sich auch die Ordnerstruktur des Anwenders befindet. Der Workbench stellt das optische Interface dar, also das, was der Nutzer zu sehen bekommt. Zur Entwicklung von grafischen Benutzeroberflächen bietet das SWT verschiedene Bausteine (Buttons, Frames, Menüs u.a.) an [FHW02]. JFace ist eine Benutzeroberflächenframework für allgemeine Aufgaben und um die Fähigkeiten der SWT-Komponenten zu erweitern [SHA04]. Die Komponente Team oder auch bekannt als Version and Configuration Management (VCM) ermöglicht eine Entwicklung im Team, d.h. es kann gleichzeitig an mehreren Projekten gearbeitet werden. Help ist die erweiterungsfähige Online-Hilfe [FHW02].

Diese Architektur beschreibt Eclipse als durch Plug-Ins erweiterbare IDE. Seit der Version 3.0 basiert die Softwarearchitektur auf dem OSGi-Framework Equinox.

Die Entscheidung den TwitterSearcher in Eclipse Version 4.3 zu entwickeln beruht auf der Aktualität der Version. Verglichen mit den Vorgängerversionen gibt es dabei einige Änderungen. Die Struktur von Eclipse-Anwendungen wird über ein logisches Modell namens Application Model beschrieben. Dieses kann während der Entwicklung und zur Laufzeit geändert und erweitert werden. Das Aussehen von Eclipse Widgets kann über externe CSS-Dateien konfiguriert werden. Das Programmiermodell basiert auf der Dependency Injection [VOG13]. Dieses Framework ermöglicht den Zugriff auf benötigte Objekte, ohne wissen zu müssen woher die jeweiligen Instanzen kommen. Die benötigten Objekte werden vom Context bezogen und aktualisiert. Vorteilhaft erweist sich, dass dadurch ein „Part weder Interfaces implementieren, noch von irgendwelchen Frameworkklassen ableiten muss“ [ECL13 Seite 16]. Das ermöglicht eine einfachere Migration von Nicht-RCP-Anwendungen zu RCP-Anwendungen [ECL13]. Es ist nicht auf Java begrenzt. Die Kompatibilitätsschicht ermöglichen es Plug-Ins, die das Eclipse 3.x Programmiermodell benutzen, unverändert in einer Eclipse 4.x Anwendung zu verwenden. EGit ist für die Versionskontrolle verantwortlich [VOG13]. Viele weitere Elemente sind bereits anhand der Abbildung 5.2 hinreichend beschrieben worden.

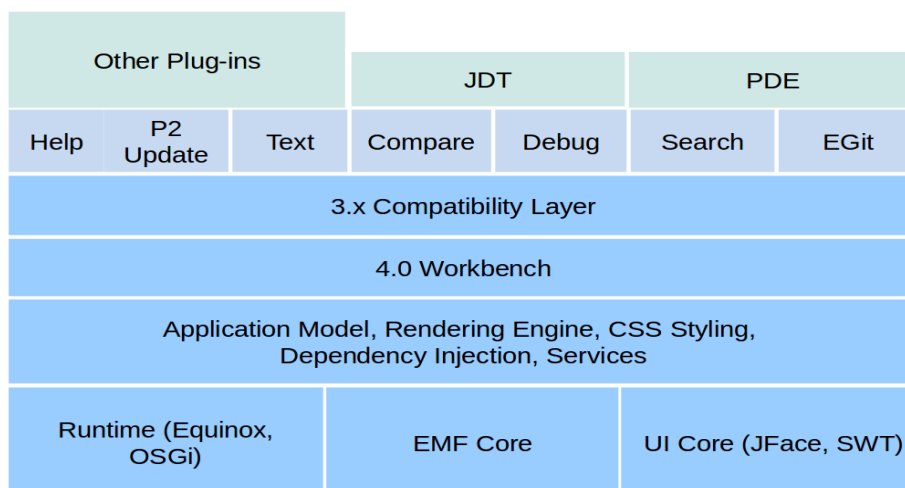


Abbildung 18 Eclipse Architektur Version 4 [VOG13]

5.1.2 OSGi-Service Plattform

„OSGi ist eine javabasierte Softwareplattform, die die dynamische Integration und das Fernmanagement von Softwarekomponenten, sogenannten Bundles und Services ermöglicht“ [WÜT08 Seite 12]. Als vorteilhaft zu erwähnen ist dabei, dass die Bundles und Services zur Laufzeit in den jeweilige Plattform installiert und gestartet beziehungsweise gestoppt und deinstalliert werden können, ohne die Plattform neu zu starten (Hot Deploement). Der ursprüngliche Einsatzort der OSGi Service Plattform waren sogenannte „*residential internet gateways*“, die in der Gebäudesystemtechnik als Verbindung zwischen hausinternen Gebäudebussen und dem Internet eingesetzt wurden [WÜT08].

Die Spezifikationen von OSGi werden von der OSGi Alliance gefördert, die sich aus verschiedenen Technologieunternehmen (z.B. Adobe, IBM, Oracle, Deutsche Telekom u.a.) zusammensetzt [OSG14].

Die Vorteile, die sich aus dieser Technologie ergeben, umfassen ein einfacheres Codemanagement (Schreiben, Testen, Wiederverwendung), die schnellere Bugerkennung, die Softwarebereitstellung ist einfacher zu handhaben und die Runtime bietet einen enormen Einblick, in das was momentan läuft [OSG14].

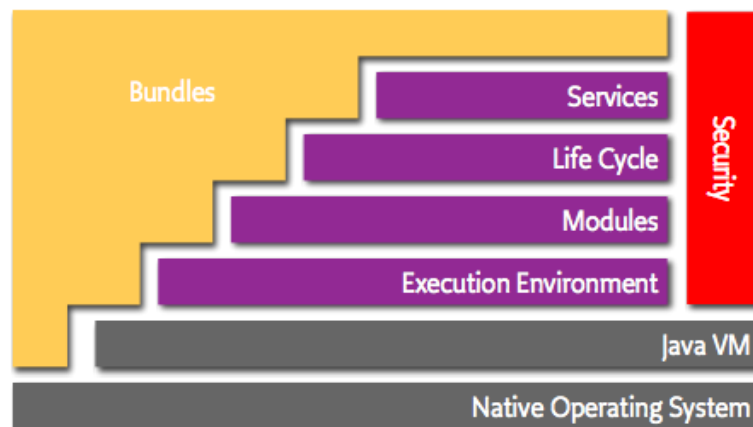


Abbildung 19 OSGi-Framework Architektur [OSG14]

Die Architektur des OSGi-Frameworks, besteht aus verschiedenen logischen Schichten (Layer) (siehe Abbildung 5.4). Das OSGi-Framework stellt einen Container dar, um Bundles und Services zur Laufzeit zu implementieren. Zudem ist es auf den unterschiedlichsten Java-Plattformen lauffähig [WÜT08].

Zu den Komponenten der OSGi-Architektur zählen die bereits erwähnten Bundles (JAR-Datei → Java Archiv), welche Komponenten von Entwicklern darstellen, die aus Klassen und Ressourcen bestehen. Diese können eigenständig im OSGi-Framework installiert und deinstalliert werden. Die Packages eines Bundles sind zunächst für andere Bundle nicht sichtbar, um dies zu ändern müssen sie durch das implementierende Bundle exportiert werden. Die Klassen und Ressourcen, die in einem Bundle enthalten sind, sind zunächst für andere Bundles nicht sichtbar. Um das zu ändern, muss das entsprechende Bundle die Packages explizit exportieren und die Packages müssen gleichzeitig von dem nutzenden Bundle importiert werden, was in einer Manifest-Datei geschieht [WÜT08].

Ebenfalls zur Entkopplung unterschiedlicher Module dienen Services. Diese verbinden Bundles auf dynamische Weise, durch das Anbieten eines Find-Bind-Systems [OSG14]. Diese Java-Objekte werden systemweit zur Verfügung gestellt, indem sie unter einem oder mehreren Interface-Namen bekannt gemacht werden. Registriert und abgemeldet werden sie in der *Service Registry*, die Bundles erlaubt den benötigten Dienst abzufragen, ohne wissen zu müssen welcher Bundle den Service anbietet [WÜT08]. Des Weiteren kann ein Bundle auch auf einen speziellen Service warten bis dieser erscheint und bekommt daraufhin einen

Call Back. Ein Bundle ist also in der Lage einen Service zu registrieren, zu bekommen und er kann hören ob ein Service bereit steht oder nicht. Beliebige viele Bundles können den Selben Servicetyp registrieren und auch bekommen (Abbildung 5.4.) [OSG14].

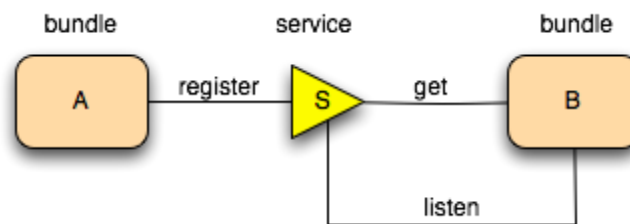


Abbildung 20 gemeinsame Nutzung eines Services von zwei Bundles [OSG14]

Wenn mehrere Bundle-Objekte unter dem Selben Interface registriert werden, muss es eine Möglichkeit geben diese zu unterscheiden. Die Lösung sind Properties. Sobald ein Service registriert wurde, besitzt er eine Reihe von Standard und benutzerdefinierten Eigenschaften. Somit ist es möglich über Properties den benötigten Service zu finden. Positiv zu erwähnen wäre noch, dass sich viele spezifische APIs mit der Service Registry besser modellieren lassen [OSG14].

Die Lebenszyklus (Life Cycle)-Schicht beschreibt die API um Bundles zu installieren, zu starten, zu stoppen, zu aktualisieren und zu deinstallieren, also die Zustände die ein Bundle während seines Lebenszyklus besitzen kann. Außerdem werden die Aktionen beschrieben, die zu Änderungen des Bundlezustandes führen [WÜT08].

Die Modul-Schicht definiert ein Bundle als grundlegende Modularisierungseinheit und wie Bundles Quellcode importieren und exportieren können. Module sind Softwareeinheiten, die technisch zusammenhängende Klassen und Ressourcen zusammenfassen und dabei die Schnittstelle und die Implementierung voneinander trennen. Dabei stellt die Modulschnittstelle die Klassen und Interfaces bereit die nach außen hin sichtbar sind, wohingegen die Implementierung nach außen nicht sichtbar ist [WÜT08].

Das Executive Environment wurde definiert um das OSGi-Framework unabhängig von einer konkreten JRE-Version zu spezifizieren – Repräsentationen von Java-Laufzeitumgebungen. Es wird festgelegt welche Klassen, Interfaces und Methoden in der jeweiligen Plattform verfügbar sind [WÜT08].

Die Security-Schicht regelt die sicherheitsrelevanten Aspekte wie z.B. Umgang mit signierten Bundles oder die Einschränkung von Ausführungsrechten von einzelnen Bundles [WÜT08].

Das javabasierte Framework Equinox stellt die OSGi-Implementierung von Eclipse dar. Mitte 2003 etablierte sich die OSGi Service Platform auf Grund der Eclipse RCP-Anwendungen um Plug-Ins während der Laufzeit zu installieren und deinstallieren ohne die Anwendung neu starten zu müssen. Wie bereits angesprochen wurde dann erstmals mit Eclipse Version 3.0

auf Eclipse Equinox umgestellt. Letztendlich besteht Equinox auch aus einer Reihe von Bundles die verschiedene OSGi-Services implementieren [WÜT08].

Mit dem Plug-In Development Enviroment (PDE) besteht eine gute Entwicklungsumgebung für Bundles. Neben den spezifischen OSGi-Anforderungen implementiert Equinox einige proprietäre Features wie die Extension Registry [WÜT08].

5.1.3 Rich Client Platform

Seit der Version 3.0 basiert Eclipse auf RCP. Rich Clients implementieren applikationsspezifische Funktionalitäten (z.B. interaktive Benutzerschnittstellen, Performance) direkt beim Client (z.B. bei Desktop-Anwendungen) [ITW14]. Dabei wurde der Begriff eingeführt, um eine Unterscheidung zu Terminal-Client-Anwendungen herzustellen [MCA06]. „Rich Clients müssen nicht bei jeder Benutzerinteraktion einen Server Request absenden und erlauben daher eine optimale Client/Server-Kommunikation“ [BOS10 Seite 8]. Einige RCP-Anwendungen sind offlinefähig, d.h., dass ihre Daten nur zu bestimmten Zeitpunkten mit dem Server synchronisiert werden [BOS10].

Die standardmäßige Rich-Client-Anwendung ist eigenständig ausführbar und verfügt über eine grafische Benutzeroberfläche (UI), die wiederum über verschiedene Steuerelemente verfügt. Rich Clients ermöglichen in einem grafischen Interface diverse Funktionen wie Drag and Drop, Multiple Fenster und Drop-Down-Menüs u.a. Rich-Client-Anwendungen basieren auf einem dynamischen Plug-In Modell [MCA06]. Ein großer Vorteil besteht darin, dass man zur Erzeugung von RCP-Anwendungen lediglich zwei Plug-Ins (org.eclipse.ui und org.eclipse.core.runtime) sowie die dazu gehörenden Voraussetzungen benötigt [VOG13].

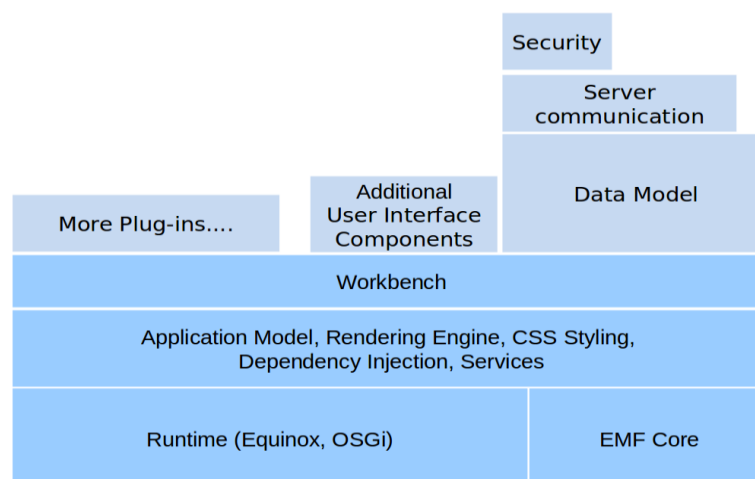


Abbildung 21 Eclipse RCP-Architektur [VOG13]

Eine Eclipse RCP-Anwendung verwendet die Basiskomponenten der Eclipse Plattform und fügt zusätzliche spezifische Komponenten der Anwendung hinzu [VOG13].

5.1.4 Eclipse Modeling Framework

Das EMF ist, wie der Name schon sagt ein Framework zur modellbasierten Softwareentwicklung und bietet dafür verschiedene Bibliotheken und Werkzeuge an. Im Daten-Modell vereinigen sich die Technologien XML, UML und Java. Diese Technologie eignet sich besonders gut bei der Umsetzung von Datenzugriffen und -haltungen [STE09].

Im Mittelpunkt der Entwicklung steht die Erzeugung eines strukturierten Modells in UML. Dieses Meta-Modell wird ecore genannt und wird durch Klassendiagramme realisiert. Diese Metaklassen besitzen Attribute, Operationen und Referenzen. Die Erzeugung kann auf verschiedene Arten erfolgen, z.B. direkt als XMI-Datei, mit dem EMF-Editor von Eclipse oder über annotierte Java-Interfaces u.a. Anhand dieses Modells wird der Java-Strukturcode, ein Editor und verschiedene Adapterklassen generiert. Die Generierung erfolgt über ein sogenanntes Generatormodell, welches die Informationen (z.B. den Pfad und Dateiinformationen) für die Codegenerierung enthält [STE09].

Der große Vorteil der sich aus dem EMF ergibt, besteht darin, dass mit minimalem Aufwand, Modelle erzeugt, und aus ihnen der entsprechende Quellcode generiert wird, ohne weitere komplexe Werkzeuge nutzen zu müssen. Damit wird eine nützliche Verbindung zwischen der Modellierung und der Programmierung geschaffen.

5.2 Programmentwicklung des TwitterSearchers

5.2.1 Motivation

Zu Beginn bestand die Idee der Kriminalisten, aufgrund der steigenden Popularität sozialer Netzwerke, ihre forensische Spurensuche, zusätzlich zu den eingeleiteten Maßnahmen, in diesen selbst auszuweiten. Da eine manuelle Untersuchung des Datenmaterials hinsichtlich Zeit, Menge und Relevanz unmöglich ist, stand die Entwicklung eines forensischen Softwaretools im Vordergrund.

Der TwitterSearcher durchsucht Tweets nach selbst definierten Schlagworten oder Nutzern und liefert eine eingeschränkte Ergebnismenge (die Gründe warum diese nicht vollständig ist, ist im Kapitel 3.2.4 hinreichend erläutert) von gefundenen Tweets mit ihren Eigenschaften. Diese werden aufgrund der Übersichtlichkeit getrennt ausgegeben. Dem Anwender steht es zusätzlich frei, in welcher Sprache er den Suchlauf startet.

Ob verwertbare Hinweise in den gefundenen Tweets enthalten sind, kann über die integrierte Textanalyse herausgefunden werden. Der TwitterSearcher schränkt die Suchmenge je nach Qualität der eingegebenen Schlagworte erheblich ein und verkürzt damit die Bearbeitungszeit durch den Kriminalisten.

5.2.2 Use Case Diagramm

Das Ziel eines jeden Softwareentwicklungsprozesses besteht darin, Software zu entwickeln, die bestimmte Anforderungen erfüllt. So steht am Anfang dieses Prozesses die Zielsetzung, also für wen soll meine Software entwickelt werden und warum. Da sich die Ziele während des Entwicklungsprozesses ändern können oder im Vorfeld nicht klar definiert wurden, kann es zu Beeinträchtigungen während der Entwicklung kommen [SCH10].

Ein Use-Case-Diagramm (Anwendungsfalldiagramm) hat die Aufgabe, eine grobe Ordnung in die detaillierten Anforderungen des Programms zu bringen. Zusätzlich soll es einen Überblick über das verschaffen, was die Software im Wesentlichen können muss. Wichtige Funktionen werden zueinander in Beziehungen gesetzt und die technische Implementierung spielt dabei keine Rolle [SCH10].

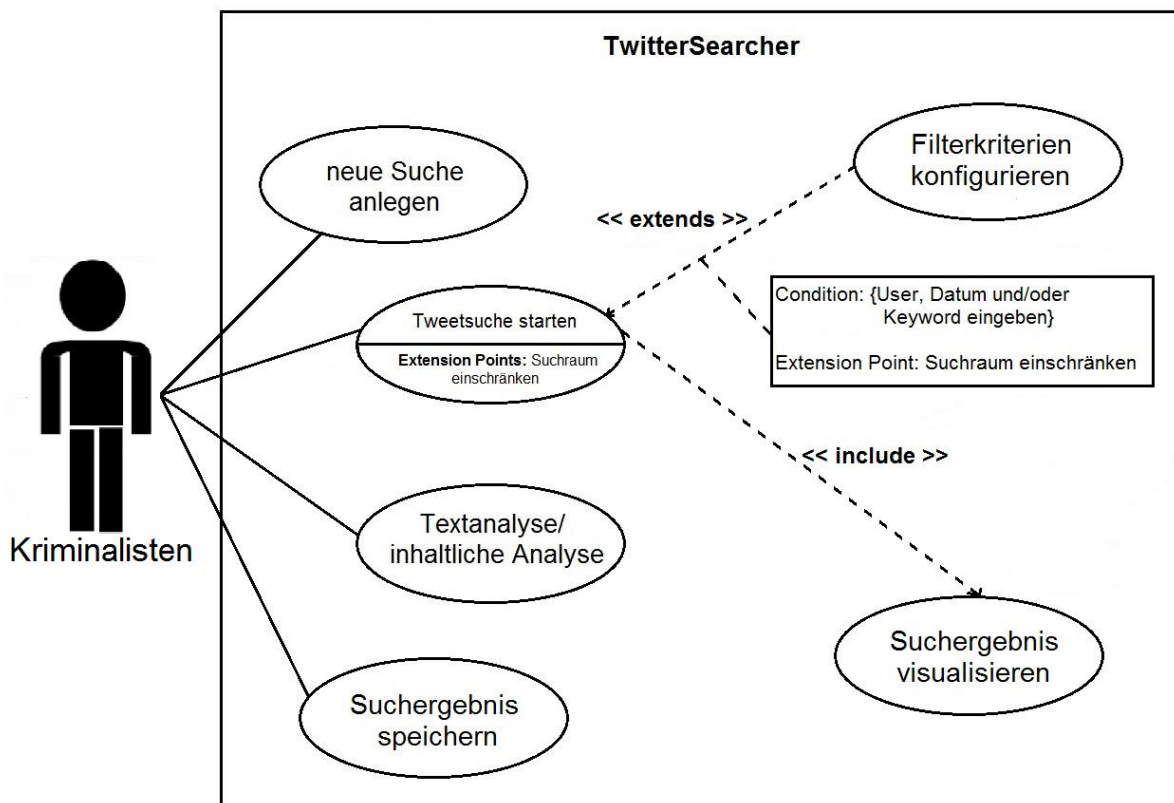


Abbildung 22 Use Case-Diagramm für den TwitterSearcher

Das dargestellte Use-Case-Diagramm zeigt den TwitterSearcher aus der Sicht des Anwenders, der in einem solchen Fall Akteur (wird immer als Strichmännchen dargestellt) genannt wird. Bezüglich des TwitterSearchers, sind die Kriminalisten die mit dem Programm arbeiten die Akteure. Sie befinden sich immer außerhalb des Systems („TwitterSearcher“), da sie nicht zum System gehören, sondern nur mit diesem interagieren. Der TwitterSearcher bietet Kriminalisten im Wesentlichen 6 Funktionen an, die als Ellipsen dargestellt sind und von denen er 4 (dargestellt mit durchgezogenen Linien) aktiv beeinflussen kann. Nachfolgend werden alle Use-Cases näher erläutert:

- **neue Suche anlegen**

Dieser Use-Case ermöglicht dem Anwender, wie es der Name schon sagt, eine neue Suchanfrage an Twitter anzulegen, indem das Programm gestartet wird. Diese Option kann beliebig oft ausgeführt werden und ist bereits vollends umgesetzt.

- **Tweetsuche starten**

Sollten sich die Anwender entschieden haben eine neue Suche anzulegen, sollte es auch das Bestreben sein, diese über einen Button in der Menüleiste zu starten. Jedoch ist dies nur möglich, wenn im Vorfeld bestimmte Filterkriterien (Use Case „Filterkriterien konfigurieren“) im Requestteil eingegeben wurden (das `<< extends >>` bedeutet, dass der zweite Use Case (Tweetsuche starten) nur ausgeführt wird, wenn der erste (Filterkriterien konfigurieren) erfüllt worden ist). Die Kriminalisten haben die Möglichkeit diese Filterkriterien hinsichtlich eines bestimmten Twitternutzers, eines Datums sowie einer White- beziehungsweise Blacklistsuche, mit selbst definierten Suchbegriffen, zu wählen und somit die gestellte Suchanfrage zu konkretisieren. Auch dieser Use Case ist bereits gänzlich umgesetzt, zeigt jedoch eine praktische Hürde auf. Da die verwendete Streaming-API von Twitter nur eine begrenzte Anzahl von Anfragen pro Stunde zulässt, kann diese Option nicht beliebig oft angewendet werden.

- **Textanalyse/inhaltliche Analyse**

Wenn Tweets hinsichtlich der gestellten Suchanfrage gefunden wurden, müssen diese nach forensisch relevanten Hinweisen (Wörtern z.B. in Form von Abkürzungen, Umgangssprache u.a., Hashtags, Bildern und Links) untersucht werden. Da die manuelle Suche auf Grund der immensen Datenmenge zeitlich nicht zu realisieren ist, wird dies softwarebasiert umgesetzt. Die Programmierung der Textanalyse ist zum jetzigen Zeitpunkt jedoch nicht umgesetzt worden.

- **Suchergebnis speichern**

Um die gerichtliche Verwertbarkeit der Daten zu gewährleisten, müssen alle Untersuchungsschritte, gesicherten Daten sowie alle anderen extrahierten Informationen dokumentiert und gespeichert werden. Jedoch ist dieser wichtige Ansatz, um vor Gericht dienliche und verwendbare Informationen zu präsentieren im aktuellen Entwicklungsstand des TwitterSearchers noch nicht realisiert worden. Was jedoch bereits umgesetzt ist, ist die Speicherung von gestellten Suchanfragen.

- **Suchergebnis visualisieren**

Zu diesem Use Case besteht keine direkte Assoziation zum Anwender, da die Darstellung, sobald eine Tweetsuche gestartet wird, vom System automatisch im Responsepart generiert wird. Das Ergebnis wird dem Anwender übersichtlich dargestellt. So wird vermerkt, wie viele und welche Tweets aufgrund der gestellten Suchanfrage gefunden wurden. Diese können einzeln angeklickt werden und der TwitterSearcher liefert danach noch tweetspezifische Informationen, wie Datum,

User, ID, Standort, die eigentliche Nachricht und eventuelle Hashtags und Links die im Tweet enthalten sind. Dieser Aspekt wurde bereits komplett programmiertechnisch umgesetzt. Jedoch sollte auch hier zu bedenken sein, dass dieses dargestellte Ergebnis nur ein Bruchteil der Suchanfrage darstellt, da die Search-API nur 1% - 40% aller Tweets auflistet (siehe Kapitel 3.2.4.1).

5.2.3 Einbinden der Twitter-API

Um Zugriff auf die Daten von Twitter zu erhalten, benötigt man eine API, die der Microbloggingdienst zur Verfügung stellt (welche verschiedenen APIs es gibt, wird im Kapitel 3.2.4 beschrieben). Für den TwitterSearcher fiel die Wahl auf die Streaming-API. Diese kann auf der Webseite <https://dev.twitter.com/docs/twitter-libraries> unter dem Punkt *Java* → *Twitter4J* kostenlos gedownloadet werden, jedoch muss im Vorfeld eine Anmeldung bei Twitter erfolgen, um die Streaming-API nutzen zu können. Auf der Homepage der Twitter Developer (<https://dev.twitter.com/docs/api/1.1>) findet man zusätzlich noch eine detaillierte Dokumentation über die Ressourcen, die durch die Streaming-API bereitgestellt werden.

Nachdem die JAR-Datei gedownloadet wurde, muss sie dem Eclipse-Project, in dem sie eingesetzt werden soll, hinzugefügt werden. Mit einem Rechtsklick auf das ausgewählte Projekt im Package Explorer, öffnet sich ein Menü in dem man „*Properties*“ auswählt. Im neuen geöffneten Fenster wählt man auf der linken Seite „*Java Build Path*“ und rechts den Reiter „*Libraries*“. Über „*Add External JARs ...*“ kann dann die Bibliothek (beim TwitterSearcher: *twitter4j-core-3.0.3.jar*) zum Projekt hinzugefügt werden und wird im Package Explorer unter den Referenced Libraries angezeigt.

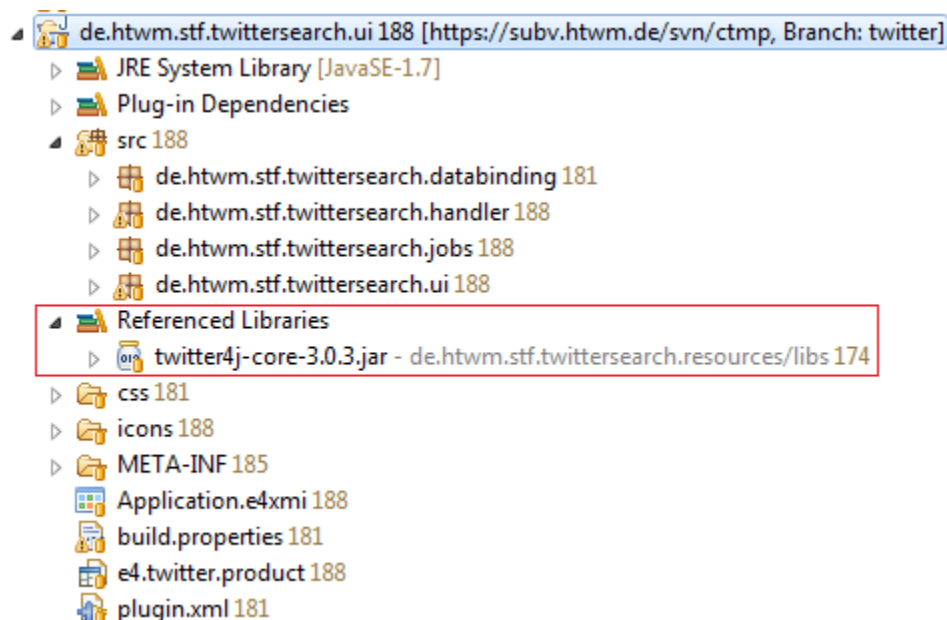


Abbildung 23 Ausschnitt aus der Ordnerstruktur des TwitterSearchers mit Referenced Libraries

Um die einzelnen Klassen der Twitter4J-Bibliothek in den konkreten Java-Klassen des Projektes nutzen zu können, bedarf es des Import-Kommandos (im nachfolgenden Beispiel, wird explizit die Klasse `TwitterExceptions` importiert): `import twitter4j.TwitterExceptions`

5.2.3.1 Entwurf der grafischen Benutzeroberfläche

5.2.3.2 Window Builder

Zur Entwicklung der grafischen Benutzeroberfläche wurde der SWT-Designer Window Builder verwendet. Um das Benutzerinterface zu verändern, hat man zum einen die Möglichkeit, dies über die Bearbeitung des Quelltextes zu tun oder unter Verwendung des integrierten Grafikeditors, der mit Parts, SWT und JFace Dialogen und JFace Wizards arbeitet. Unterstützt werden Eclipse 3.x und Eclipse 4 RCP-Anwendungen.

Die Installation kann am einfachsten mit Update Manager von Eclipse (Menüleiste: *Help* → *Update Manager*) vorgenommen werden. Nach der Eingabe der Eclipse Kepler Update Website (<http://download.eclipse.org/releases/kepler>) findet man in der Auflistung, unter „General Purpose Tools“ den SWT Designer, womit die Installation fortgesetzt wird und nach wenigen weiteren Schritten ihren Abschluss findet.

Nach einer erfolgreichen Installation wird eine Toolpalette zur Verfügung gestellt, die diverse Komponenten wie z.B. Layouts, Tabellen, Buttons, Textfelder, Labels, JFace-Elemente, Menü-Elemente und speziell für den TwitterSearcher Nebula-Elemente bereithält. Über Drag and Drop können diese Komponenten aus der Palette auf der zu entwickelnden Benutzeroberfläche platziert werden.

5.2.3.3 Grafische Benutzeroberfläche des TwitterSearcher

Innerhalb der `Application.e4xmi` wird die ursprüngliche Struktur der RCP-Anwendung, anhand des Application Modells beschrieben. Nachfolgend wird der strukturelle Aufbau der kompletten Oberfläche beschrieben.

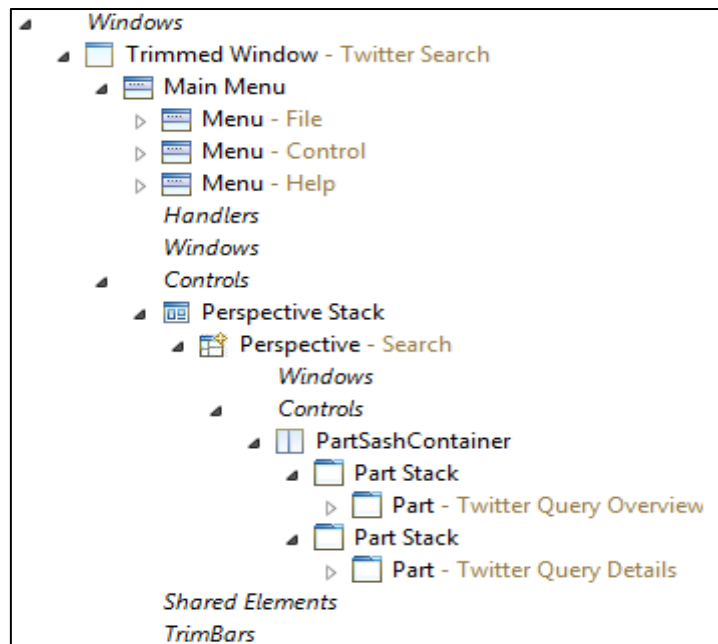


Abbildung 24 Ausschnitt der Windowstruktur aus der Application.e4xmi

Unter dem Strukturpunkt „*Windows*“ wird ein „*Trimmed Window*“ hinzugefügt. Im Gegensatz zum normalen „*Window*“, welches alternativ als Child von „*Windows*“ gewählt werden kann, verfügt das Trimmed Window über sogenannte TrimBars, was z.B. Toolbars, Statusleisten oder Stacks einschließt, die beliebig im Fenster angeordnet werden können [ECP10]. Nach einem Klick auf das Trimmed Window, besteht die Möglichkeit, in einem rechts erscheinenden Fenster, verschiedene Parameter, wie die ID (muss zur eindeutigen Erkennung immer vergeben werden), Abgrenzungen (Höhe, Breite u.a.) und Labels u.a. anzugeben. Das Trimmed Window verfügt über ein Main Menü, welches im Grunde ebenfalls ein Child darstellt und wiederum über 3 Childs (File, Control und Help) verfügt. Durch einen Klick auf die Menüpunkte, können rechts ebenfalls die Parameter abgerufen werden. Innerhalb dieser Menüpunkte sind verschiedene Interaktionen möglich.

Der TwitterSearcher verfügt also über ein Main Menü mit 3 Menüpunkten. Im File-Menü befinden sich weiterhin die Punkte **Save** (noch nicht umgesetzt), zum Speichern der Suchanfrage und **Exit** (verfügt zusätzlich über ein Icon) um das Programm zu verlassen. Das Control-Menü verfügt über den **Start**-Menüpunkt, über den die Anwendung gestartet werden kann. Zusätzlich verfügt er über ein Icon (grüner Kreis mit weißem Pfeil → Anlehnung an das Run von Eclipse um Anwendungen zu starten), welches unter dem Parameter Icon URI eingestellt werden kann. Der Menüpunkt Help, beinhaltet den **About**-Menüpunkt, der Auskunft über die Anwendung gibt. All diese Interaktionen werden über Handlerklassen realisiert, über die später noch Ausführungen gemacht werden.

In den „*Controls*“ befindet sich der „*Perspective Stack*“, der Auskunft über die Betrachtungsweise des TwitterSearchers gibt. Der PartSashContainer teilt die Oberfläche horizontal in zwei Bereiche (Details Part und Query Overview Part), die sich innerhalb eines PartStacks (Layout Container) befinden und in den nächsten Kapiteln näher beschrieben werden.

5.2.3.3.1 Details Part

Der Detailspart beinhaltet alle Elemente um eine neue Suchanfrage mit bestimmten Filterkriterien anzulegen und diese visualisieren zu lassen.

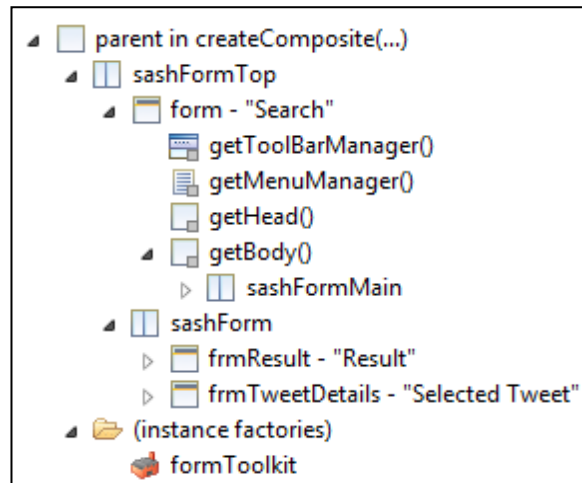


Abbildung 25 Windowstruktur des Details Part

Die Struktur beginnt mit dem Punkt „*parent in createComposite*“ in der die Zusammensetzung der einzelnen Form erzeugt wird. Diese Form ist wiederum horizontal in zwei Abschnitte geteilt. Zum einen befindet im oberen Teil der Anwendung die Form „Search“ mit all ihren Elementen. Der Body dieses Abschnittes ist ebenfalls durch eine SashForm in 3 Sektionen (Whitelist, Blacklist, OtherFilters) unterteilt, die anhand der später aufgezeigten Benutzeroberfläche deutlich werden. Diese Sektionen können in der Toolpalette unter Forms API gefunden werden, welche nachträglich installiert werden muss. Des Weiteren ist der untere Abschnitt in gleicher Weise horizontal durch eine SashForm in zwei Formen (frmResult und frmTweetDetails) unterteilt die ebenfalls anhand der Benutzeroberfläche mit ihren Funktionen verdeutlicht werden.

Durch einen Klick auf das jeweilige Element, können von diesem in einem weiteren Fenster die Parameter bearbeitet und eingesehen werden. So ist ersichtlich, dass der Detailspart auf einem Filllayout beruht, welches alle Komponenten der Reihe nach anordnet und den gesamten Raum des Containers damit füllt. Die einzelnen Sektionen verfügen dann wiederum über ein Formlayout

The screenshot shows a web interface for a Twitter search tool. It is divided into three main sections under the heading 'Search'.

- Whitelist:** Labeled 'Hier die gesuchten Begriffe eingeben.' (Here enter the terms you are searching for). It features a text input field with a small 'i' icon, a green plus button, and a red minus button. Below the input is a large empty rectangular box.
- Blacklist:** Labeled 'Hier die nicht-gesuchten Begriffe eingeben.' (Here enter the terms you do not want to search for). It has the same layout as the Whitelist section with an input field, plus/minus buttons, and a large empty box.
- Other Filter Options:** Labeled 'Metadaten editieren.' (Edit metadata). It contains:
 - A 'User (confirm with Enter)' section with a text input field and a small 'i' icon.
 - A 'Date' section with a text input field showing '10.04.2014' and a calendar icon.
 - A 'Language' section with a dropdown menu.

Abbildung 26 Searchform des Detail Parts

Die Searchform hat die Aufgabe, eine geeignete Suchanfrage an Twitter mit verschiedenen Filterkriterien zu erstellen. Besonders gut ersichtlich ist dabei die beschriebene Sektioneneinteilung, die im Vorfeld innerhalb der Struktur angelegt worden ist. Bezüglich der Abbildung 5.10 werden anschließend die Elemente der Searchform näher erläutert.

Die linke Sektion trägt, wie bereits erwähnt, den Namen Whitelist. Sie ermöglicht dem Anwender die Eingabe von selbstdefinierten Keywords nach denen er explizit in den Tweets suchen möchte. Über ein Eingabefeld können diese Keywords, durch den grünen Plusbutton der Auswahlliste hinzugefügt werden. Der rote Minusbutton löscht nicht mehr benötigte oder falsche Keywords aus der Auswahlliste. Alle eingegebenen Keywords werden für eine spätere Verwendung gespeichert und können jederzeit wieder abgerufen werden. Der Tooltip links neben dem Eingabefeld weist nachdrücklich nochmals auf diese Funktion hin.

Die Blacklist-Sektion verfügt über die selben Elemente zur Definition von Filterkriterien wie die Whitelist. Jedoch dient die Blacklist dem Zweck, dass hier Keywords eingegeben werden können, nach denen in den Tweets nicht gesucht werden soll. Beide Listenbereiche können parallel oder einzeln genutzt werden (jedoch ist eine alleinige Blacklistsuche nicht unbedingt sinnvoll). Je konkretere Keywords der Anwender an das System stellt, umso genauer ist letztendlich auch das Suchergebnis.

Other Filter Options ermöglicht eine Suche nach bestimmten Metadaten eines Tweets. So besteht die Möglichkeit explizit nach Tweets eines bestimmten Users zu suchen. In diesem Fall wird auf den Plus- und Minusbutton verzichtet, da die Suche jeweils nur nach einem User möglich ist und die Bestätigung über die Enter-Taste erfolgt. Komplettiert werden die Metadaten von einer Datumssuche, die über das Widget DateTime realisiert worden ist. Dieses zeigt beim Start des Programms immer das aktuelle Datum an und kann über den integrierten Kalender auf das gewünschte Datum eingestellt werden. Hier besteht die Möglichkeit nach Tweets zu suchen die an einem bestimmten Tag gesendet worden sind. Des Weiteren besteht noch die Option, Tweets hinsichtlich einer bestimmten Sprache zu suchen. Über die Combobox Language können Tweets in den Sprachen Deutsch, Englisch, Spanisch, Französisch und Italienisch gesucht werden. Diese Auswahl ist jederzeit erweiterbar. Auch hier besteht die Möglichkeit es untereinander parallel, einzeln zu nutzen oder in Verbindung mit der White- beziehungsweise Blacklistsuche.

The image shows a user interface for a Twitter search application. It is divided into two main vertical panels. The left panel, titled 'Result', shows '0 Tweet(s) found.' and a large empty rectangular area below it. The right panel, titled 'Selected Tweet', contains metadata fields: 'Date:', 'ID:', 'User:' (with a circular profile picture icon), and 'Place: not available'. Below these is a section titled 'Message' with a right-pointing arrow. At the bottom of the right panel is a section titled 'New Section' with a downward-pointing arrow, containing two empty rectangular boxes side-by-side.

Abbildung 27 Result- und Selected Tweetform des Detail Parts

Der untere Teil des Details Part dient der Visualisierung der Suchergebnisse und ist dabei durch eine SashForm in zwei Forms unterteilt.

Nachdem die Suchkriterien in der oberen Form konfiguriert worden sind, kann die Anfrage über das Startmenü in der Menüleiste gestartet werden. Die Ergebnisse, die mit der Anfrage übereinstimmen werden in der linken Resultliste dargestellt. Zusätzlich erfolgt eine Zählung der gefundenen Tweets (es sollte nochmals erwähnt sein, dass die Anzahl der gefundenen Tweets nur ein Bruchteil der eigentlichen Suchmenge ist). Möchte der Anwender konkretere Information zu den gefundenen Tweets haben, so reicht ein Klick auf den jeweiligen Tweet und diese werden in der Selected Tweetform aufgezeigt.

In der Selected Tweetform werden die angewählten Tweets hinsichtlich ihrer Metadaten übersichtlich dargestellt. Im oberen Bereich erfolgt die Darstellung des Sendedatums, des Verfassers (User), der ID des Tweets sowie des Sendeortes (noch nicht umgesetzt), der über die Geolocalization-Funktion festgestellt werden kann. Die Darstellung des Verfassers, bietet noch einen weiteren Vorteil, da dieser verlinkbar abgebildet wird und somit zu allen Tweets des Users führt.

Darunter befinden sich wiederum zwei weitere Sektionen, wobei die untere von beiden durch eine SashForm horizontal in zwei Listen unterteilt ist. In der Message-Sektion wird der komplette Tweettext angezeigt. Die darunter befindlichen Listen dienen dazu im Tweet enthaltene Links (linke Liste) und Hashtags (rechte Liste) abzubilden.

Alle Sektionen des Details Part verfügen über die Collapsible-Funktion, die es gestattet nicht benötigte Sektionen einzuklappen und so für mehr Übersichtlichkeit zu sorgen.

5.2.3.3.2 Query Overview Part

Die komplette linke Seite des TwitterSearchers wird dabei vom Twitter Query Overview Part eingenommen (in Abbildung 5.12, rot gekennzeichnet), der eine Liste aller bisher gestellten Suchanfragen speichert. Durch das Anwählen eines Eintrages wird die entsprechende Suchanfrage mit den dazugehörigen Ergebnissen wieder aufgerufen.

Sowohl der Details Part als auch der Query Overview Part lassen sich einzeln auf der grafischen Benutzeroberfläche darstellen, abhängig von den Interaktionen die der Anwender gegenwärtig mit dem TwitterSearcher ausführen möchte.

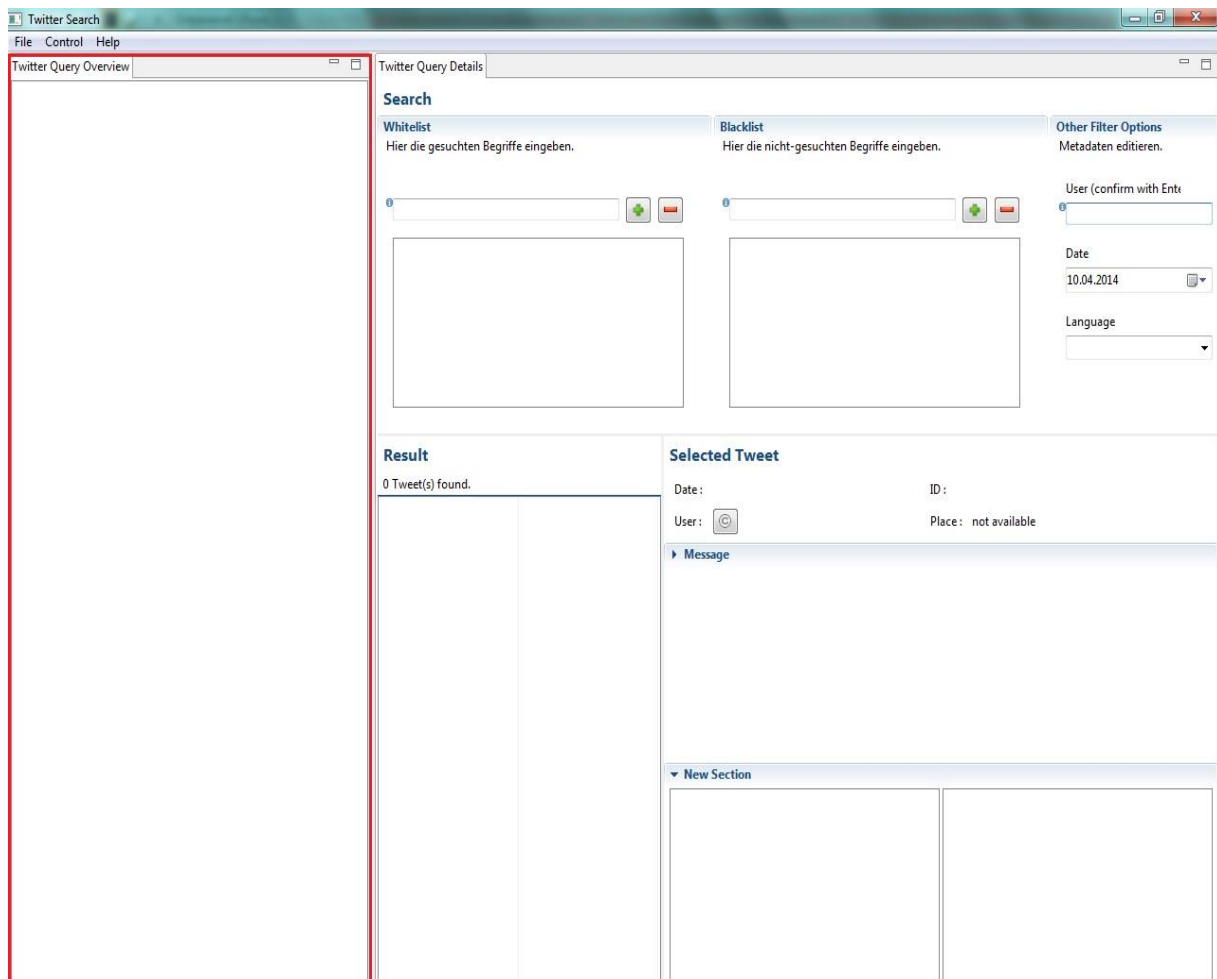


Abbildung 28 grafische Benutzeroberfläche des TwitterSearchers mit gekennzeichnetem Query Overview Part

5.2.4 Datenmodell

Das EMF-Datenmodell des TwitterSearchers dient dazu, die Metadatenstrings, die über die Streaming-API von Twitter bereitgestellt werden, zu speichern und daraus die Strings für die jeweiligen Listen und Textfelder auf der grafischen Benutzeroberfläche zu konstruieren. Generiert wurde das nachfolgende Datenmodell aus dem Ecore. Nachfolgendes Datenmodell wurde in der grafischen Modellierungssprache UML erzeugt.

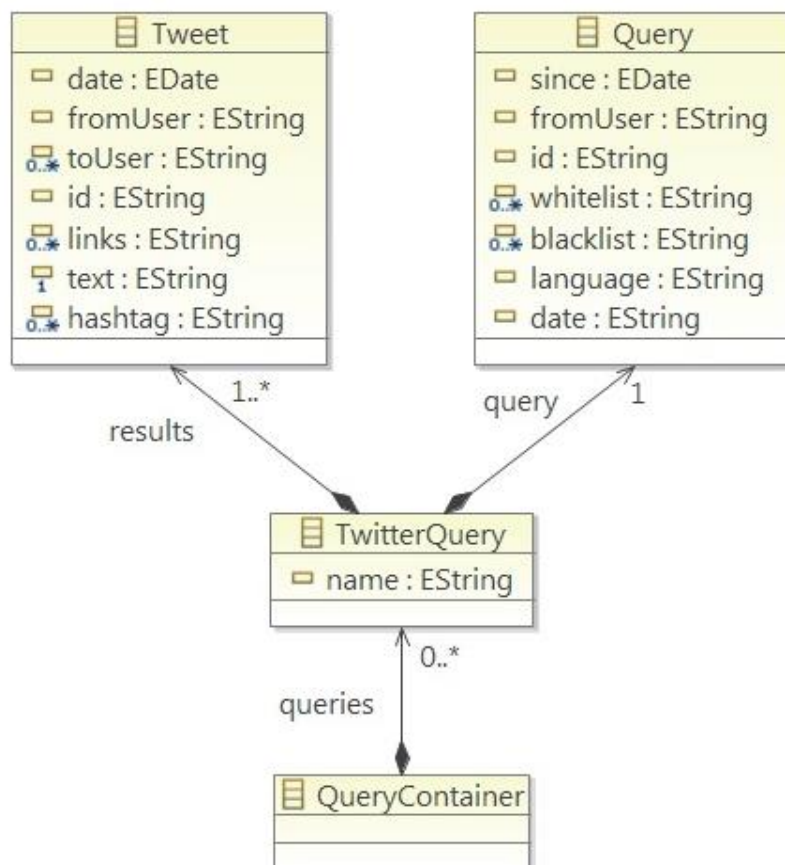


Abbildung 29 EMF-Datenmodell des TwitterSearchers

Das dargestellte Datenmodell zeigt 4 EClasses. Die EClass **Tweet** besitzt dabei mehrere EAttribute (`date`, `fromUser`, `toUser`, `id`, `links`, `text`, `hashtag`), die die Bestandteile eines Tweets repräsentieren. Jedem EAttribute wurde dabei ein bestimmter Datentyp zugewiesen. Entweder ein `EString`, für eine Zeichenkette, oder `EDate` für ein Datum. Diese EAttribute werden später für die Ausgabe im Result und Selected Tweet Part verwendet. Für die EAttribute `toUser`, `links`, und `hashtag` wurde ein Lower Bound und ein Upper Bound (`0..*`) hinzugefügt, da von diesen Attributen mehrere innerhalb eines Tweets vorkommen können beziehungsweise beim Attribut `toUser` der Tweet an mehrere User geschickt werden kann.

Weiterhin existiert noch die EClass **Query**, die ebenfalls mehrere EAttribute (`since`, `fromUser`, `id`, `whitelist`, `blacklist`, `language`, `date`) mit den dazugehörigen Datentypen besitzt und für die Zusammensetzung der Suchanfrage zuständig ist. Für die EAttribute `whitelist` und `blacklist` wurden ebenfalls Bounds (`0..*`) festgelegt, da die Möglichkeit besteht ganz auf die White-

beziehungsweise Blacklistsuche zu verzichten (0) oder bekanntermaßen die Option vorhanden ist, so viele Keywords wie möglich in die erwähnten Listen einzutragen.

Die EClass `TwitterQuery` mit ihrem EAttribute `name` dient dazu, einen Ergebnisstring zu speichern. Zum einen besteht eine Assoziation (Komposition) zur Query namens `query`, mit einer festgelegten Bound von 1, was bedeutet, dass sie nur aus genau einer Query besteht. Des Weiteren besteht noch eine zusätzliche Assoziation (Komposition) zu `Tweet`, mit einer Begrenzung von `1...*` und bedeutet, dass sie zusätzlich zu der einen Query aus mindestens einem Tweet bestehen muss.

Eine Komposition beschreibt eine Sonderform des Assoziations-Modellelementes. Gekennzeichnet durch ein schwarzes Karo am Ende der Assoziation. „Bei einer Komposition kann ein Teil immer nur in genau einem zusammengesetzten Objekt enthalten sein, und die Lebensdauer des zusammengesetzten Objekts entspricht immer der Lebensdauer seiner Komponenten. Das zusammengesetzte Objekt wird hier als Kompositum bezeichnet“ [LAH09 Seite 129].

In der EClass `QueryContainer` werden alle gestellten Suchanfragen gespeichert und später im Query Overview Part dargestellt. Die bestehende Assoziation (Komposition) namens `queries` zur EClass `TwitterQuery` bedeutet, der `QueryContainer` eine endlose Anzahl von Twitter Queries aufnehmen kann.

Basierend auf diesem Datenmodell wird dann über *File → New → Eclipse Modeling Framework → EMF Generator Model* ein neues Generatormodell erzeugt, indem man das erstellte Ecore-Modell auswählt und dieses Generatormodell dann in der Auszeichnungssprache XML dargestellt [VOG15].

Auf Grundlage dieser beiden Modelle (Ecore und Genmodel) wird im nächsten Schritt der Java-Code erzeugt. Nach einem Rechtsklick und der Auswahl *Generate Model Code* wird die Java Implementation des Modells erzeugt. Dieser Code besteht aus einem Modell (beim `TwitterSearcher` „`twitterquery`“), einer Modell Implementation („`twitterquery.impl`“), welche eine Implementierung des Interfaces innerhalb des Modells darstellt und einer `AdapterFactory` („`twitterquery.util`“) [VOG15]. In der `TwitterQueryAdapterFactory` sind Methoden enthalten, die dem Erzeugen aller definierten Objekte dienen. Das Modell („`twitterquery.impl`“) enthält GET- und SET-Methoden für jedes Attribut (z.B. in der `Tweet.java`: `getDate()`, `getID()`, `setDate()`, `setID()` u.a.). Jede generierte Methode wird dabei mit *@generated* gekennzeichnet. Jedes erzeugte Interface extends das `EObject` Interface: ***public interface Tweet extends EObject***, was die Grundlage jeder EMF-Klasse [VOG15].

5.2.6 Services

Zur Entwicklung von Eclipse RCP-Anwendungen and Plug Ins werden sogenannte Services verwendet. Services sind Softwarekomponenten, die entweder auf einem Interface oder

einer Klasse beruhen und eine bestimmte Funktionalität bereitstellen. Die erzeugten Klassen basieren auf dem erstellten Datenmodell [VOG14].

Der dem TwitterSearcher zugrunde liegende Service trägt den Namen *TwitterSearchModelService*. In der Regel dient er dazu, Elemente im Modell zu suchen, neue Elemente zu erzeugen oder zu entfernen und er gewährt Zugriff auf das Datenmodell während der Laufzeit, um die Daten zu verändern [VOG14]. Der *ModelService* innerhalb des *TwitterSearchers* dient u.a. dazu die gestellten Suchanfragen in den *QueryContainer* zu schreiben, wo sie dann im *Query Overview Part* auf der grafischen Benutzeroberfläche dargestellt werden. Des Weiteren sollen alle Suchanfragen mit ihren Ergebnissen gespeichert werden und jede neue Suchanfrage dem *QueryContainer* im Datenmodell hinzugefügt werden, welche dann ebenfalls im *Query Overview Part* dargestellt werden sollen.

5.2.7 Handlerklassen

Um die verschiedenen Aktionen im Hauptmenü ausführen zu können, wird auf verschiedene Handlerklassen zurückgegriffen. Diese werden im Fall des *TwitterSearchers* durch einen Mausklick ausgelöst. Jedem Menüpunkt werden als Childs sogenannte *HandledMenuItems* übergeben, in denen dann im Feld „Command“ die ID des dazugehörigen Kommandos („abstraktes Kommando ohne definierte Semantik“ [DAU05 Seite 38]) übergeben wird. Dies erfolgt in der Windowstruktur durch eine ID innerhalb der *Application.e4xmi*.

- *InfoHandler*: dient zur Anzeige eines selbst definierten Infotextes
- *StartHandler*: dient der Ausführung des Programmes nachdem konkrete Filterkriterien eingegeben wurde
- *ExitHandler*: schließt des Programm
- *DynamicPartHandler*: dient der Erstellung eines neuen Parts innerhalb des aktuellen Fensters

Zusätzlich muss später noch eine Handlerklasse, bezüglich der Speicherung der sichergestellten Daten erzeugt werden.

6 Schlussbetrachtungen

Zusammenfassend muss erwähnt werden, dass der TwitterSearcher mit seinen bisher umgesetzten Funktionen die Erwartungen und den Nutzen, aufgrund der steigenden Popularität sozialer Netzwerke erfüllt. Die noch ausstehenden Funktionen beeinträchtigen jedoch eine Übergabe an die Polizeidirektion Chemnitz und sollten die nächsten Schritte sein, um die programmtechnische Umsetzung abzuschließen. Besonders die Speicherung der sichergestellten Daten sollte dabei oberste Priorität haben, um diese gerichtlich verwertbar zu machen.

Ferner könnten die Bestrebungen um Programmerweiterungen dahingehen, dass unter der Verwendung eines Visualisierungstoolkits, wie z.B. Prefuse, Userbeziehungen grafisch dargestellt werden.

So erfreulich die Programmentwicklung verlaufen ist, so ärgerlich fällt das Fazit über die erhaltenen Suchergebnisse aus. Wenn man bedenkt, dass Twitter in Deutschland bei Weitem (noch) nicht die Nutzerzahlen wie Facebook hat und der TwitterSearcher (wie der Name schon sagt) ausschließlich Nachrichten von Twitter nach gestellten Anfragen durchsucht, so sollten die erhaltenen Ergebnisse nicht überbewertet werden. Erschwerend kommt hinzu, dass die verwendete Streaming-API des Microbloggingdienstes nur einen Bruchteil aller täglich verfassten Tweets liefert, was die Suche nach lokal relevanten Begebenheiten weiter einschränkt.

Hinsichtlich dieses Problems gilt zu überlegen, ob die Nutzung der kostenpflichtigen Firehose-API, die alle Daten liefert, sinnvoll wäre. Hier stellt sich jedoch die Kostenfrage. Um diese Überlegung überhaupt anzustreben, müsste sich der TwitterSearcher für die Kriminalisten, hinsichtlich der gefundenen Hinweise und den damit eventuell verbundenen Fortschritten bei der Aufklärung von Straftaten im Vorfeld mehrfach ausgezeichnet haben. Auch Facebook besitzt eine kostenpflichtige API, die das Suchspektrum weiter ausdehnen könnte.

Der ursprüngliche Wunsch der Kriminalisten, lokale Netzwerke nach Hinweisen durchsuchen zu können, wird nicht realisierbar sein, da diese ihre API nicht frei zugänglich machen. Jedoch würden die Suchergebnisse, mit mehr integrierten sozialen Netzwerken innerhalb des forensischen Analysewerkzeuges, mehr Gehalt und Aussagekraft liefern. So ähneln sie vielmehr einer Stichprobe und können kaum verwertbare Sachen enthalten.

Literaturverzeichnis

- [BEC06] Beck, Klaus: Computervermittelte Kommunikation im Internet : Lehr- und Handbücher der Kommunikationswissenschaft. – München : Oldenbourg, 2006
- [BKA12] Bundeskriminalamt [Hrsg.]: Cybercrime Bundeslagenbild 2012. - Wiesbaden, 2012 URL: http://www.bka.de/DE/Publikationen/JahresberichteUndLagebilder/Cybercrime/cybercrime__node.html?__nnn=true
- [BOD06] Bode, Elisa ; Greiner, Claudia ; Fritzsche, Monika ; König, Stefanie: Topologien. URL: <http://home.arcor.de/team2/topologien.htm>, 2006
- [BOS10] Bosch, Tobias [Hrsg.]: Rich Clients – In oder out? URL: consulting.com/fileadmin/redaktion/veroeffentlichungen/pdf/javaspektrum-2010-06-rich-clients-bosch-sicher.pdf?ref=ocweb&web_c=rss&web_k=Presse_aktuell, 2010
- [BRI13] BrightPlanet Corporation. Twitter Firehose vs. Twitter API: What's the difference and why should you care? URL: <http://www.brightplanet.com/2013/06/twitter-firehose-vs-twitter-api-whats-the-difference-and-why-should-you-care/>, 2013 (Stand 10.02.2014)
- [BUG14] Buggisch, Christian: 11 Millionen oder eine halbe Million? Wie viele Twitter-Nutzer gibt es in Deutschland? URL: <http://buggisch.wordpress.com/2014/02/04/wie-viele-twitter-nutzer-gibt-es-in-deutschland/>, (Stand 10.02.2014)
- [COM11] IDG Business Media GmbH. Hülsbörner, Simone [Hrsg.]: Die 12 besten Twitter-Alternativen. URL: <http://www.computerwoche.de/a/die-12-besten-twitter-alternativen,1912070>, 2011
- [DAT14] DATASIFT. URL: <http://datasift.com/> (Stand 01.13.2014)
- [DAU05] Daum, Berthold: Rich-Client Entwicklung mit Eclipse 3.1: Anwendungen entwickeln mit der Rich Client Platform. – 1.Auflage – Heidelberg : dpunkt Verlag, 2005
- [DEW11] Dewald, Andreas ; Freiling Felix C.: Forensische Informatik. – 1.Auflage – Erlangen : Books on Demand GmbH Norderstedt, 2011
- [DÖR06] Döring, Nicola: Computervermittelte Kommunikation. – Hagen : FernUniversität, Fakultät für Kultur- und Sozialwissenschaften, 2006

- [ECL13] Ford-Reitz, Zeb (Hrsg.): Dependency Injection (DI). In: eclipse-Magazin. 2013, Nr.5, Seite 16. URL:
http://www.bredex.de/files/BredexDateien/publications/ECM_5_13_Ford-Reitz_Migration.pdf
- [ECL14] The Eclipse Foundation: Platform architecture: URL:
<http://help.eclipse.org/indigo/index.jsp?topic=%2Forg.eclipse.platform.doc.isv%2Fguide%2Farch.htm>
- [ECP10] Eclipsepedia: E4/UI/Modeled UI. URL:
http://wiki.eclipse.org/E4/UI/Modeled_UI#Window, 2010
- [EDU05] EDUCATION GROUP GmbH. e-LISA academy: Beispiele und Aufbau von Blogs. URL:
<http://www.e-lisa-academy.at/?PHPSESSIONID=&design=elisawp&url=community&cid=7470&modul=10&folder=52203&>, 2005
- [FAZ12] Frankfurter Allgemeine Zeitung GmbH. Hansen-Nootbaar, Anna [Hrsg.]: Mit Twitter und Facebook Straftaten verhindern. URL:
<http://www.faz.net/aktuell/gesellschaft/oeffentliche-sicherheit-mit-twitter-und-facebook-straf-taten-verhindern-11848541.html>, 2012 (Stand: 08.08.2012)
- [FWW02] Schramm, Daniel (FH Wedel): Die Eclipse-Architektur. URL: *<http://www.fh-wedel.de/~si/seminare/ws02/Ausarbeitung/7.eclipse/eclipse2.htm>*, Wedel, 2002
- [FSD10] Petrovic, Sasa ; Osborne, Miles ; Lavrenko, Victor: Streaming First Story Detection with application on Twitter. – 2010. Edinburgh, University, School of Informatics. URL: *www.aclweb.org/anthology/N10-1021.pdf*
- [FRA12] Fraas, Claudia ; Meier, Stefan ; Pentzold, Christian: Online-Kommunikation: Grundlagen, Praxisfelder und Methoden. – München : Oldenbourg Verlag, 2012
- [GNI14] Gnip, Inc. URL: *<http://gnip.com/>* (Stand 01.03.2014)
- [HOL13] Holl, Alfred: Knowledge Management – Wissensmanagement. – 2013. Nürnberg, University of Applied Science, URL:
http://www.informatik.fh-nuernberg.de/professors/Holl/Personal/B_Slides_C_KM.pdf
- [ITW14] DATACOM Buchverlag GmbH (Hrsg.): Rich Client. URL:
<http://www.itwissen.info/definition/lexikon/Rich-Client-rich-client.html>

- [KIO14] Kioskea.net: Funktionsweise von Bluetooth. URL: <http://de.kioskea.net/contents/605-funktionsweise-von-bluetooth>, (Stand April 2014)
- [KUR08] Kurose, James F. ; Ross, Keith W.: Computernetzwerke: Der Top-Down-Ansatz. – 4.Auflage – München : Pearson Studium, 2008
- [LAH09] Lahres, Bernhard ; Rayman, Gregor: Objektorientierte Programmierung Das umfassende Handbuch. – 2.Auflage. – Bonn : Galileo Press, 2009
- [LAN13] Landesakademie für Fortbildung und Personalentwicklung an Schulen: E-Mail: IMAP statt POP. URL: http://lehrerfortbildung-bw.de/werkstatt/internet/pop_imap/ (Stand Januar 2014), 2013
- [LIN13] Linten, Martin ; Schemberg, Axel ; Surendorf, Kai: PC-Netzwerke: Das umfassende Handbuch. - 6.Auflage – Bonn : Galileo Press, 2013
- [MAN13] Manca, Matteo ; Boratto, Ludovico ; Carta, Salvatore: Producing Friend Recommendations in a Social Bookmarking System by Mining User Content – 2013. Cagliari, Universita, URL: http://www.thinkmind.org/index.php?view=article&articleid=immm_2013_3_20_20041
- [MCA06] McAffer, Jeff ; Lemieux, Jean-Michel: eclipse Rich Client Platform - Designing, Coding, and Packaging Java Applications. – Boston : Pearson Education, Inc, 2006
- [MIC14] Microsoft (Hrsg.): Was ist der Unterschied zwischen Hubs, Switches, Routern und Zugriffspunkten. URL: <http://windows.microsoft.com/de-de/windows/hubs-switches-routers-access-points-differ#1TC=windows-7>, (Stand Januar 2014)
- [MOO14] Moos, Roland: Funktionsprinzip Internet Telefonie. URL: <http://www.voip-information.de/voip-voice-over-ip.php>, (April 2014)
- [MÜN07] Münz, Stefan SELFHTML e.V.: Foren und Boards. URL: <http://aktuell.de.selfhtml.org/artikel/gedanken/foren-boards/>, (Stand Januar 2014), 2007

- [N24] N24 Gesellschaft für Nachrichten und Zeitgeschehen mbH: Das sind Deutschlands beliebteste soziale Netzwerke. URL: <http://www.n24.de/n24/Nachrichten/Netzwelt/d/4217156/das-sind-deutschlands-beliebteste-soziale-netzwerke.html>, 2014
- [NET] Netzthemen.de: Die Entstehung eines neuen Medien-Genres. URL: <http://www.netzthemen.de/sonnabend-weblogs/2-2-1-die-entstehung-eines-neuen-medien-genres>, (Stand Januar 2014)
- [NÖL05] Nölle, Jochen: Voice over IP: Grundlagen, Protokolle, Migration – Berlin, Offenbach : VDE-Verlag, 2005
- [OSG14] OSGi Alliance: The OSGi Architecture. URL: <http://www.osgi.org/Technology/WhatIsOSGi>
- [PKS12] Bundesministerium des Inneren (Hrsg.): Polizeiliche Kriminalstatistik 2012. – Berlin, 2012. URL: https://www.bmi.bund.de/SharedDocs/Downloads/DE/Broschueren/2013/PKS2012.pdf?__blob=publicationFile
- [ROB12] Roberts, Kirk ; Roach, Michael ; Johnson, Joseph ; Guthrie, Josh ; Harabagiu, Sanda: EmpaTweet: Annotating and Detecting Emotions on Twitter. – 2012, Dallas, University, Human Language Technology Research Institute. URL: http://www.lrec-conf.org/proceedings/lrec2012/pdf/201_Paper.pdf
- [SCH07] Schmidt, Jan: Entwicklung des Internets und ausgewählter Anwendungen. URL: <https://www.flickr.com/photos/42154456@N00/2119923557/>, 2007
- [SCH10] Schäling, Boris: Der moderne Softwareentwicklungsprozess mit UML. URL: <http://www.highscore.de/uml/usecasesdiagramm.html>, 2010
- [SCH13] Scheffler, Tatjana: Computerlinguistische Analyse von Twitterdaten. – 2013. Potsdam, Universität, Department Linguistik. URL: <http://www.ling.uni-potsdam.de/~scheffler/teaching/2013twitter.html>
- [SHA04] Shavor. Sherry ; D'Anjou, Jim ; Fairbrother, Scott ; Kehn, Dan ; Kellerman, John ; McCarthy, Pat: Eclipse Anwendungen und Plug-Ins mit Java entwickeln. – München: Addison-Wesley Verlag, 2004
- [SIM10] Simon, Nicole ; Bernhardt, Nikolaus: Twitter: Mit 140 Zeichen zum Web 2.0 – 2.Auflagen – München, Open Source Press, 2010

- [SPR] Springer Gabler Verlag (Hrsg.), Gabler Wirtschaftslexikon, Stichwort: Schwarzes Brett. URL
<http://wirtschaftslexikon.gabler.de/Archiv/75923/schwarzes-brett-v9.html>,
(Stand 15.01.2014)
- [STA14] Statista GmbH: Anzahl der aktiven Nutzer von Facebook in Deutschland von Januar 2010 bis Januar 2014 (in Millionen). URL:
<http://de.statista.com/statistik/daten/studie/70189/umfrage/nutzer-von-facebook-in-deutschland-seit-2009/>, 2014
- [STE03] Stetina, Birgit: Aufbau einer E-Mail Adresse – Universität Wien. Zentraler Informatik Dienst. URL:
<http://homepage.univie.ac.at/birgit.stetina/flskryspin/archiv/emailauf.PDF>,
(Stand 15.01.2014), 2003
- [STE09] Steinberg, Dave ; Budinsky, Frank ; Paternostro, Marcelo ; Merks, Ed: EMF Eclipse Modeling Framework. – 2.Auflage – Boston : Pearson Education, Inc, 2009
- [TAN03] Tanenbaum, Andrew S: Computernetzwerke. – 4.Auflage – München : Pearson Studium, 2003
- [TIA13] Tian, Tian ; Chun, Ae Soon: Social Networked-based Entity Extraction for People Ontology. – 2013. Manhattan, College, Department of Computer Science. New York, University, College of Staten Island URL:
http://www.thinkmind.org/index.php?view=article&articleid=immm_2013_3_30_20096
- [TIK] TIKRIS Tim Kröner Internet Services. Was ist VoIP bzw. Voice over IP? URL:
<http://www.voip-information.de/voip-voice-over-ip.php> (Stand 16.01.2014)
- [TIL10] TILL.DE GmbH: Weltweite Online-Studie: E-Mails werden von Sozialen Netzwerken abgelöst. URL:
<http://blog.till.de/allgemein/weltweite-online-studie-e-mails-werden-von-sozialen-netzwerken-abgelost/>, (Stand 15.01.2014), 2010
- [TWI12] Twitter Developers. The Streaming APIs. URL:
<https://dev.twitter.com/docs/api/streaming>, 2012 (Stand 05.02.2014)
- [TWI13] Twitter Developers. Using the Twitter Search API. URL:
<https://dev.twitter.com/docs/using-search>, 2013 (Stand 05.02.2014)

- [TWI14] Twitter. Deaktivieren deines Account. URL: <https://support.twitter.com/groups/51-me/topics/205-account-settings/articles/106854-deaktivieren-deines-accounts#>, (Stand 20.03.2014)
- [UVT14] Universität Trier: Was ist Computerlinguistik? Fach Computerlinguistik und Digital Humanities. URL: <https://www.uni-trier.de/index.php?id=20241>
- [VOG13] Vogella. Vogel, Lars (Hrsg.): Building Eclipse RCP applications based on Eclipse 4. URL: http://www.vogella.com/tutorials/EclipseRCP/article.html#e4overview_eclipse4, 2013
- [VOG14] Vogella. Vogel, Lars (Hrsg.): Eclipse platform services – Tutorial. URL: <http://www.vogella.com/tutorials/Eclipse4Services/article.html>, 2013
- [VOG15] Vogella. Vogel, Lars (Hrsg.): Eclipse Modeling Framework (EMF) – Tutorial. URL: <http://www.vogella.com/tutorials/EclipseEMF/article.html>, 2013
- [WDR09] Westdeutscher Rundfunk [Hrsg.] ; Südwestrundfunk [Hrsg.]: Planet Wissen: Geschichte des Computers. URL: http://www.planet-wissen.de/wissen_interaktiv/html-versionen/geschichte_des_computers/ (Stand 30.01.14), 2009
- [WDW10] Schön, Friederike: 10 Wahrheiten über Facebook. In: Welt der Wunder. 2010, Nr.6, Seite 51.
- [WDW12] Ramos, Nuno ; Cholet, J.: Tatort Facebook – Wie Verbrecher das soziale Netzwerk erobern. In: Welt der Wunder. 2012, Nr.5, Seite 91.
- [WIK14] Wikipedia Foundation Inc. Tf-idf-Maß. URL: <http://de.wikipedia.org/wiki/Tf-idf-Maß>, (Stand 21.03.2014)
- [WIK141] Wikipedia Foundation Inc.: Entität (Informatik) URL: [http://de.wikipedia.org/wiki/Entität_\(Informatik\)](http://de.wikipedia.org/wiki/Entität_(Informatik)), (Stand März.2014)
- [WIK142] Wikipedia Foundation Inc.: Ontologie (Informatik) URL: [http://de.wikipedia.org/wiki/Ontologie_\(Informatik\)#Bestandteile](http://de.wikipedia.org/wiki/Ontologie_(Informatik)#Bestandteile), (Stand März.2014)
- [WIK143] Wikipedia Foundation Inc.: E-Mail. URL: <http://de.wikipedia.org/wiki/E-Mail> (Stand April 2014)

- [WIK144] Wikipedia Foundation Inc.: Eclipse (IDE). URL:
http://de.wikipedia.org/wiki/Eclipse_%28IDE%29, (Stand April 2014)
- [WLA14] Schröck, Thorsten: Heimnetzwerk – Einstieg in die Heimvernetzung
Teil 1 URL:
<http://schroeck-multimedia.de/heimnetzwerk-einstieg-in-die-heimvernetzung-teil-1/>, 2014
- [WUN13] Wunderwelt Wissen: Entertain your brain: Technik schnell erklärt – Wie sterbe ich im Internet. September 2013, Nr.9, Seite 86.
- [WÜT08] Wütherich, Gerd ; Hartmann, Nils ; Kolb, Bernd ; Lübken, Matthias: Die OSGi Service Platform: Eine Einführung mit Eclipse Equinox. – 1.Auflage – Heidelberg : dpunkt.verlag GmbH, 2008
- [YEE13] Yeebase Media GmbH. Tantau, Björn [Hrsg.]: Blaue Vögel haben's schwer: Warum Twitter in Deutschland nicht durchstartet. URL:
<http://t3n.de/news/twitter-in-deutschland-481600/>, 2013
- [ZAR12] Zarella, Dan [Hrsg.]: Eigenschaften eines perfekten Tweets! URL:
<http://www.seo-united.de/blog/social-media/eigenschaften-eines-perfekten-tweets.htm>, 2012

Anhang

Selbstständigkeitserklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe.

Stellen, die wörtlich oder sinngemäß aus Quellen entnommen wurden, sind als solche kenntlich gemacht.

Diese Arbeit wurde in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegt.

Groitzsch, den 16.04.2014